

Mon Apr 3 08:24:16 2000

us-08-852-495c-1_copy_1_30000.rge

Page 2

```

KEYWORDS      histone 1; promoter.
SOURCE        Homo sapiens DNA.
ORGANISM      Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
REFERENCE     Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS       Gallinari,P., La Bella,F. and Heintz,N.
TITLE         Characterization and purification of H1RP2, a novel CCAAT-binding
              protein that interacts with a histone H1 subtype-specific consensus
              element
JOURNAL       Mol. Cell. Biol. 9 (4), 1566-1575 (1989)
MEDLINE       89261784
COMMENT       On Jul 23, 1994 this sequence version replaced gi:3414458.
FEATURES
  source      1..85
              /organism="Homo sapiens"
              /db_xref="taxon:9606"
              /cell_line="HeLa"
              /bound_molecule="SP1"
  CAAT_signal 53..70
              /gene="HIS1"
  gene         53..85
              /gene="HIS1"
  TATA_signal  77..85
              /gene="HIS1"
  BASE_COUNT   26 a          27 c          18 g          14 t
  ORIGIN
Query Match           0.2%; Score 46; DB 10; Length 85;
Best Local Similarity 100.0%; Pred. No. 9,6e-12;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 26024 GTGCCTAGGTGATGCACCATCACCACGGCGCCCTATATTAAAG 26069
Db 40 GTGCCTAGGTGATGCACCATCACCACGGCGCCCTATATTAAAG 85
|||||
|||

RESULT 2
LOCUS       AF042300                73 bp    DNA             STS            26-JUN-1999
DEFINITION Homo sapiens chromosome 6 map 6p21.3, sequence tagged site.
ACCESSION   AF042300
VERSION     AF042300.1 GI:4191287
KEYWORDS    STS.
SOURCE      human.
            Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
REFERENCE   Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS     Lin,L., Jin,L., Kimura,A., Carrington,M. and Mignot,E.
TITLE       DQ microsatellite association studies in three ethnic groups
JOURNAL     Tissue Antigens 50 (5), 507-520 (1997)
REFERENCE   2 (bases 1 to 73)
            /bases 1 to 73
            Lin,L., Jin,L., Voros,A., Underhill,P. and Mignot,E.
            Microsatellite single nucleotide polymorphisms in the HLA-DQ region
            Tissue Antigens 52 (1), 9-18 (1998)
JOURNAL     98378271
            3 (bases 1 to 73)
            Lin,L., Jin,L., Lin,X., Voros,A., Underhill,P. and Mignot,E.
            Direct Submision
            Submitted (12-JAN-1998) Psychiatry and Behavioral Sciences,
            Stanford University, Lab Surge Building, 1201 Welch Road, Room
            P126, Palo Alto, CA 94304, USA
FEATURES
  source      1
              /organism="Homo sapiens"
              /db_xref="taxon:9606"
              /chromosome="6"
              /map="6p21.3"
repeat_region 1..73

```

```

/!note="microstellite"
/rpt_unit=ca
/db_type=tandem
BASE COUNT 32 a 29 c 3 g 9 t 9 t
ORIGIN

Query Match 0.1% Score 40; DB 13; Length 73;
Best Local Similarity 100.0%; Pred. No. 1.le-08;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CACACACACACACACACACACACACACACACACAAATG 40
Db 17 CACACACACACACACACACACACACACACACACAAATG 56

RESULT 3
AF042301 75 bp DNA STS 26-JAN-1999
LOCUS Homo sapiens Chromosome 6 map 6p21.3, sequence tagged site.
DEFINITION AF042301
ACCESSION AF042301.1 GI:4191288
VERSION STS.
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 75)
AUTHORS Lin,L., Jin,L., Kimura,A., Carrington,M. and Mignot,E.
TITLE DQ microsatellite association studies in three ethnic groups
JOURNAL Tissue Antigens 50 (5), 507-520 (1997)
MEDLINE 98049194
REFERENCE 2 (bases 1 to 75)
AUTHORS Lin,L., Jin,L., Lin,X., Voros,A., Underhill,P. and Mignot,E.
TITLE Tissue Antigens 52 (1), 9-18 (1998)
JOURNAL 98378271
REFERENCE 3 (bases 1 to 75)
AUTHORS Lin,L., Jin,L., Lin,X., Voros,A., Underhill,P. and Mignot,E.
TITLE Direct Submission
JOURNAL Submitted (12-JAN-1998) Psychiatry and Behavioral Sciences,
Stanford University, 14b Surge Building, 1201 Welch Road, Room
126, Palo Alto, CA 94304, USA
LOCATION/Qualifiers

FEATURES
source
1..75
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="6"
/map="6p21.3"
1..75
/note="microsatellite"
/rpt_unit=ca
/db_type=tandem
33 a 30 c 3 g 9 t
ORIGIN

Query Match 0.1% Score 40; DB 13; Length 75;
Best Local Similarity 100.0%; Pred. No. 1.le-08;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CACACACACACACACACACACACACACACACACAAATG 40
Db 19 CACACACACACACACACACACACACACACACACAAATG 58

RESULT 4
AF042302 77 bp DNA STS 26-JAN-1999
LOCUS Homo sapiens chromosome 6 map 6p21.3, sequence tagged site.
DEFINITION AF042302
ACCESSION AF042302.1 GI:4191289
VERSION STS.
KEYWORDS
```


Mon Apr 3 08:24:16 2000

us-08-852-495c-1_copy_1_30000.rge

Page 4

```

VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
    AU025275.1 GI:4515198
    STS.
    Rattus norvegicus DNA, clone:OT48.13/788f08.
    Rattus norvegicus
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
    Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
    1 (sites)
    Watanabe,T.K., Hishiyaki,H., Kanemoto,N., Miyakita-Mizoguchi,A.,
    Oga,K., Okuno,S., Ono,T., Tsuji,A., Hayashi,H., Adachi,M.,
    Yamashita,Y., Iriye,Y., Takahashi,E., Takagi,T., Nakamura,Y. and
    Iatanihara,Y.
    The large-scale mapping of rat microsatellite markers
    Unpublished (1998)
    2 (bases 1 to 97)
    Watanabe,T.K.
    Direct Submission
    Submitted (24-JUL-1998) to the DDBJ/EMBL/GenBank databases. Takeshi
    K. Watanabe, Otsuka Pharmaceutical Co., Ltd., Otsuka GEN Research
    Institute, 463-10, Kagasuno, Kawachi-cho, Tokushima, Tokushima
    771-0192, Japan (E-mail:watanabe@otsuka.genome.ad.jp,
    Tel:81-886-65-2888, Fax:81-886-37-1035)
    Location/Qualifiers
        1..97
        /organism="Rattus norvegicus"
        /db_xref="taxon:10116"
        /clone="OT48.13/788f08"
        /note="OT48.13/788f08F-5'-TCCAGGCCCTAGTATCACTTAA-3',
        OT48.13/788f08R-5'-CCGCGAGCATTAATTAAGCC-3'"
        16 a
        12 c
        30 g
        39 t

```

Query Match	0.1%;	Score 39;	DB 13;	Length 97;
Best Local Similarity	100.0%;	Pred. No. 3.5e+08;		
Matches	39;	Conservative	0;	Mismatches 0; Indels 0; Gaps
Oy	1	CACACACACACACACACACACACACACACACACAAAT	39	
Db	62	CACACACACACACACACACACACACACACACACAAAT	24	
RESULT	8			
LOCUS	131337			
DEFINITION	Sequence 249 from patent US 5582979.	PAT	06-FEB-1997	
ACCESSION	131337			
VERSION	131337.1	GI:1822128		
KEYWORDS				
SOURCE	Unknown.			
ORGANISM	Unknown.			
REFERENCE	1 (bases 1 to 44)			
AUTHORS	Weber, J.L.			
TITLE	Length polymorphisms in (dc-da).sub.n.(dc-dr).sub.n sequences and			
JOURNAL	method of using the same			
FEATURES	Patent: US 5582979 A 249 10-DEC-1996;			
	Location/Qualifiers			
source	1..44			
	/organism="unknown"			
BASE COUNT	23 a 21 c 0 g 0 t			
ORIGIN				

```

Query Match: Score 38; DB 5; Length 44;
Best Local Similarity 100.0%; Pred. No. 1.2e-07;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

RESULT 9

	LOCUS	131413				PAT	06-FEB-1997
	DEFINITION	Sequence 325 from patent US 5582979.					
	ACCESSION	131413					
	VERSION	131413.1					
	KEYWORDS	GI:1822204					
	SOURCE	Unknown.					
	ORGANISM	Unclassified.					
	REFERENCE	1 (bases 1 to 51)					
	AUTHORS	Weber J.L.					
	TITLE	Length polymorphisms in (dc-da).sub.n.(dg-dt).sub.n sequences and method of using the same					
JOURNAL	Patent:	US 5582979-A 325 10-DEC-1996;					
FEATURES	Location/Qualifiers						
source	1..51						
	/organism="unknown"						
BASE COUNT	27 a	24 c	0 g	0 t			
ORIGIN							
Query Match		0.1%; Score 38; DB 5; Length 51;					
Best Local Similarity	100.0%; Pred. No. 1.2e-07;						
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;							
OY	1 CACACACACACACACACACACACACACACACACAA 38						
DG	4 CACACACACACACACACACACACACACACACAAA 41						

RESULT	10				
MMND58/c					
LOCUS					
DEFINITION					
ACCESSION					
VERSION					
KEYWORDS					
SOURCE					
ORGANISM					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
DEPT. SURGERY, LEVEL 6, JOHN RADCLIFFE HOSP., HEADINGTON OXFORD, UK					
FEATURES					
source					
base					
satellite					
19 a					
8 c					
29 g					
31 t					

	Query Match:	0.1%;	Score 38;	DB 12;	Length 87;
	Best Local Similarity:	100.0%;	Pred. No. 1.1e-07;		
Matches	38;	Conservative	0;	Mismatches	0;
				Indels	0;
G?	1 CACACACACACACACACACACACACACACAA	38			
Db	62 CACACACACACACACACACACACACACAAA	25			

RESULT 11

HS2492G9
LOCUS HS2492G9 100 bp DNA STS 28-NOV-1994
DEFINITION H. sapiens (H2525) DNA segment containing (CA) repeat; clone
AFM492G9; single read, sequence tagged site.
ACCESSION 217138
VERSION 217138.1 GI:23763
KEYWORDS STS; CA repeat; dinucleotide repeat; GT repeat; microsatellite DNA;
microsatellite marker; repeat polymorphism.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 100)
Wesselsbach, J.
TITLE Direct Submission
JOURNAL Submitted (05-OCT-1992) Genethon, B.P. 60, 91002 Evry Cedex France.
AUTHORS E-mail: Jean.Wesselsbach@genethon.fr
2 (bases 1 to 100)
Wesselsbach, J., Gyapay, G., Dib, C., Vignal, A., Morissette, J.,
Millesseau, P., Vaysseix, G. and Lathrop, M.
A second-generation linkage map of the human genome
Nature 359 (6398), 794-801 (1992)
3 (bases 1 to 100)
Gyapay, G., Morissette, J., Vignal, A., Dib, C., Fizames, C.,
Millesseau, P., Marc, S., Bernardi, G., Lathrop, M. and Wesselsbach, J.
The 1993-94 Genethon human genetic linkage map
Nat. Genet. 7 (2 Spec No), 246-339 (1994)
95004593
COMMENT cloning vector is M13mp18;
full automatic.
FEATURES
source 1.100
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/cell_line="CEPH 134702"
/clone_1lb="genomic DNA"
BASE COUNT 35 a 31 c 10 g 24 t
ORIGIN

Query Match 0.1%; Score 38; DB 13; Length 100;
Best Local Similarity 100.0%; Pred. No. 1.1e-07;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CACACACACACACACACACACACACACACACAA 38
Db 28 CACACACACACACACACACACACACACACACAA 65
RESULT 12
MMVIMV11/c 102 bp DNA ROD 30-AUG-1996
LOCUS MMVIMV11/c 102 bp DNA ROD 30-AUG-1996
DEFINITION M.musculus DNA for vimentin-binding fragment VII.
ACCESSION X89128
VERSION X89128.1 GI:872135
KEYWORDS house mouse.
SOURCE Mus musculus
ORGANISM Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 102)
Wang, X., Tolstosov, G., Shoeman, R.L. and Traub, P.
Selective binding of specific mouse genomic DNA fragments by mouse
vimentin filaments in vitro
DNA Cell Biol. 15 (3), 209-225 (1996)
96226400
JOURNAL DNA Cell Biol. 15 (3), 209-225 (1996)
MEDLINE 96226400
REFERENCE Shoeman, R.L.
AUTHORS Direct Submission
JOURNAL Submitted (23-JUN-1995) Shoeman R. L., Max-Planck-Institute fuer
Zellbiologie, Rosenhof, Ladenburg, Germany, D-66526

FEATURES
source Location/Qualifiers
1..102
/organism="Mus musculus"
/db_xref="taxon:10090"
/cell_line="Ehrlich ascites tumor cell"
1..102
/standard_name="fragment VII in citation #1"
/citation=[1]
/bound_moiety="vimentin"
/function="sequence bound by vimentin filaments in vitro"
/evidence="experimental"
14..55
/citation=[1]
/rpt_family="d(CA/GT)n dinucleotide repeats"
/rpt_unit="14..15
/function="recombination, potentially to form non-B DNA"
/rpt_type="direct"
BASE COUNT 16 a 17 c 32 g 37 t
ORIGIN

Query Match 0.1%; Score 38; DB 12; Length 102;
Best Local Similarity 100.0%; Pred. No. 1.1e-07;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CACACACACACACACACACACACACACACACAA 38
Db 48 CACACACACACACACACACACACACACACACAA 11
RESULT 13
AU046973 104 bp DNA STS 22-APR-1999
LOCUS AU046973/c
DEFINITION Rattus norvegicus, OTSUKA clone, 109c04, microsatellite sequence,
sequence tagged site.
ACCESSION AU046973
VERSION AU046973.1 GI:4631608
KEYWORDS STS.
SOURCE Rattus norvegicus (strain: Brown Norway) liver hepatocyte DNA,
clone:109c04.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE 1 (sites)
Watanabe, T.K., Hishigaki, H., Okuno, S., Mizoguchi, A., Oga, K.,
Tsuji, A., Ono, T., Yamashiki, Y., Kanemoto, N., Takahashi, E., Irie, Y.,
Nakamura, Y., Takagi, Y. and Tanigami, A.
The large-scale mapping of rat microsatellite markers
Unpublished (1998)
2 (bases 1 to 104)
Watanabe, T.K.
Direct Submission
JOURNAL Submitted (11-DEC-1998) to the DDBJ/EMBL/GenBank databases. Takeshi
K Watanabe, Otsuka GEN Research Institute, Otsuka Pharmaceutical
Co., Ltd; 463-10, Kagasuno, Kawasumi-cho, Tokushima, Tokushima
771-0192, Japan (E-mail: watanabe@otsuka.gr.jp, Tel: +81-886-65-2888,
Fax: +81-886-37-1035)
FEATURES
source Location/Qualifiers
1..104
/organism="Rattus norvegicus"
/strain="Brown Norway"
/db_xref="taxon:10116"
/cell_type="hepatocyte"
/clone="109c04"
/tissue_type="liver"
/note="109c04F-5'-CCCTTCCTATTTC-3',
109c04R-5'-ATTAAAGTATAGGAATACACAT-3'."
BASE COUNT 15 a 14 c 21 g 54 t
ORIGIN

Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 ACACACACACACACACACACACACAAATGA 41
|||||
Db 73 ACACACACACACACACACACACAAATGA 36
|||||

RESULT 14
AF042298 69 bp DNA STS 26-JAN-1999
LOCUS Homo sapiens chromosome 6 map 6p21.3, sequence tagged site.
DEFINITION AF042298
ACCESSION AF042298.1 GI:4191285
VERSION STS.
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
AUTHORS Lin, L., Jin, L., Kimura, A., Carrington, M. and Mignot, E.
TITLE DQ microsatellite association studies in three ethnic groups
JOURNAL Tissue Antigens 50 (5), 507-520 (1997)
MEDLINE 98049194

REFERENCE
AUTHORS Lin, L., Jin, L., Voros, A., Underhill, P. and Mignot, E.
TITLE Microsatellite single nucleotide polymorphisms in the HLA-DQ region
JOURNAL Tissue Antigens 52 (1), 9-18 (1998)
MEDLINE 98378271

REFERENCE
AUTHORS Lin, L., Jin, L., Lin, X., Voros, A., Underhill, P. and Mignot, E.
TITLE Direct submission
JOURNAL Submitted (12-JAN-1998) Psychiatry and Behavioral Sciences,
Stanford University, Lab Surge Building, 1201 Welch Road, Room
P126, Palo Alto, CA 94304, USA
Location/Qualifiers

FEATURES
source
1..69
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="6"
/map="6p21.3"
/note="microsatellite"
repeat_region
1..69
/rpt_unit=ca
/rpt_type=tandem
BASE COUNT 30 a 27 c 9 g 9 t
ORIGIN

Query Match 0.1%; Score 37; DB 13; Length 69;
Best Local Similarity 100.0%; Pred. No. 3.7e-07;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 ACACACACACACACACACACACAAATG 40
|||||
Db 16 ACACACACACACACACACACACAAATG 52
|||||

RESULT 15
AF042305 83 bp DNA STS 26-JAN-1999
LOCUS Homo sapiens chromosome 6 map 6p21.3, sequence tagged site.
DEFINITION AF042305
ACCESSION AF042305.1 GI:4191292
VERSION STS.
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
AUTHORS Lin, L., Jin, L., Kimura, A., Carrington, M. and Mignot, E.
TITLE DQ microsatellite association studies in three ethnic groups
JOURNAL Tissue Antigens 50 (5), 507-520 (1997)
MEDLINE 98049194

REFERENCE 2 (bases 1 to 83)
AUTHORS Lin, L., Jin, L., Lin, X., Voros, A., Underhill, P. and Mignot, E.
TITLE Microsatellite single nucleotide polymorphisms in the HLA-DQ region
JOURNAL Tissue Antigens 52 (1), 9-18 (1998)
MEDLINE 98378271

REFERENCE
AUTHORS Lin, L., Jin, L., Lin, X., Voros, A., Underhill, P. and Mignot, E.
TITLE Direct submission
JOURNAL Submitted (12-JAN-1998) Psychiatry and Behavioral Sciences,
Stanford University, Lab Surge Building, 1201 Welch Road, Room
P126, Palo Alto, CA 94304, USA
Location/Qualifiers

FEATURES
source
1..83
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="6"
/map="6p21.3"
/note="microsatellite"
repeat_region
1..83
/rpt_unit=ca
/rpt_type=tandem
BASE COUNT 37 a 34 c 3 g 9 t
ORIGIN

Query Match 0.1%; Score 37; DB 13; Length 83;
Best Local Similarity 100.0%; Pred. No. 3.7e-07;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CACACACACACACACACACACACACACAA 37
|||||
Db 19 CACACACACACACACACACACACACACAA 55
|||||

Search completed: March 29, 2000, 10:09:37
Job time: 24326 sec

Mon Apr 3 08:24:16 2000

us-08-852-495c-1_copy_1_30000.rge

Page 7



•
•
•



GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 29, 2000, 03:50:02 ; Search time 599.56 Seconds

(without alignments)
12518.785 Million cell updates/sec

Title: US-08-852-495c-1_COPY_1_30000
Perfect score: 30000

Sequence: 1 CACACACACACACACACACA.....AGCCATACAGAGTGGCCCC 30000

Scoring table: OLIGO_NUC
Gapop 60.0, Gapext 60.0

Searched: 311585 seqs, 125096042 residues

Word size: 0

Total number of hits satisfying chosen parameters: 433264

Minimum DB seq length: 8

Maximum DB seq length: 105

Post-processing: Listing first 45 summaries

Database: N_Geneseq_36:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	38	0.1	44	1 T65763	Repeat sequence fr
2	38	0.1	51	1 T65773	Repeat sequence fr
3	37	0.1	60	1 Q33930	Microsatellite seq
4	37	0.1	70	1 Q33849	Microsatellite seq
5	36	0.1	36	1 Q33828	Microsatellite seq
6	36	0.1	36	1 Q33882	Microsatellite seq
7	36	0.1	36	1 Q33953	Microsatellite seq
8	36	0.1	37	1 Q34041	Microsatellite seq
9	36	0.1	37	1 Q34178	Microsatellite seq
10	36	0.1	37	1 Q33900	Microsatellite seq
11	36	0.1	37	1 T65732	Repeat sequence fr
12	36	0.1	38	1 T66048	(dc-da)n.(dg-dt)n
13	36	0.1	38	1 T65750	Repeat sequence fr
14	36	0.1	39	1 Q33648	Microsatellite seq
15	36	0.1	39	1 Q33807	Microsatellite seq
16	36	0.1	39	1 Q33825	Microsatellite seq
17	36	0.1	39	1 T65731	Repeat sequence fr
18	36	0.1	40	1 Q34091	Sequence of a micr
19	36	0.1	40	1 T65725	Repeat sequence fr
20	36	0.1	40	1 T66051	(dc-da)n.(dg-dt)n
21	36	0.1	41	1 T66054	(dc-da)n.(dg-dt)n
22	36	0.1	41	1 T65745	Repeat sequence fr
23	36	0.1	41	1 T65710	Repeat sequence fr
24	36	0.1	41	1 T65758	Repeat sequence fr
25	36	0.1	42	1 Q33770	Microsatellite seq
26	36	0.1	42	1 T65797	Repeat sequence fr
27	36	0.1	42	1 T65757	Repeat sequence fr
28	36	0.1	43	1 T65794	Repeat sequence fr
29	36	0.1	44	1 Q34113	Sequence of a micr
30	36	0.1	44	1 Q33636	Microsatellite seq
31	36	0.1	44	1 Q33983	Microsatellite seq
32	36	0.1	44	1 T65749	Repeat sequence fr
33	36	0.1	44	1 T65761	Repeat sequence fr

C	34	36	0.1	45	1 Q33968	Microsatellite seq
C	35	36	0.1	45	1 Q33915	Microsatellite seq
C	36	36	0.1	46	1 Q33840	Microsatellite seq
C	37	36	0.1	46	1 Q33939	Microsatellite seq
C	38	36	0.1	46	1 T65719	Repeat sequence fr
C	39	36	0.1	46	1 T65709	Repeat sequence fr
C	40	36	0.1	46	1 T65756	Repeat sequence fr
C	41	36	0.1	46	1 T65780	Repeat sequence fr
C	42	36	0.1	47	1 Q33834	Microsatellite seq
C	43	36	0.1	47	1 T65716	Repeat sequence fr
C	44	36	0.1	47	1 T65713	Repeat sequence fr
C	45	36	0.1	47	1 T65755	Repeat sequence fr

ALIGNMENTS

RESULT	1	
ID	T65763	
AC	T65763; standard; DNA; 44 BP.	
DT	17-JUN-1997 (first entry)	
DE	Repeat sequence from polymorphic marker clone Mf669.	
KW	Polymorphism; repeat sequence; genetic marker; primer; amplification;	
KW	PCR; polymerase chain reaction; paternity; maternity; human; pedigree;	
KW	linkage analysis; genetic disease; animal; plant; breeding; locus;	
KW	hybridisation; chromosome; ds.	
OS	Homo sapiens.	
PN	US5582979-A.	
PD	10-DEC-1996.	
PF	21-APR-1989; 341562.	
PR	21-APR-1989; US-341564.	
PR	05-SEP-1991; US-754351.	
PR	04-APR-1994; US-222177.	
PA	(MARS-) MARSHFIELD CLINIC.	
PI	Weber JL.	
DR	WPI: 97-042299/04.	
PT	Detection of polymorphic genetic markers of the form	
PS	(dc-da)n.(dg-dt)n using novel nucleic acid mols. as primers	
CC	The invention relates to the isolation of polymorphic repeat sequences	
CC	having the sequence (dc-da)n.(dg-dt)n which can be used as genetic	
CC	markers. Primers based on these sequences can be used to detect these	
CC	repeats, especially for use in e.g. paternity or maternity testing,	
CC	human genetic analysis such as linkage analysis of genetic disease,	
CC	commercial animal or plant breeding or pedigree analysis. Clones	
CC	containing the repeat sequences were isolated by hybridisation of	
CC	chromosome-specific phage libraries with a synthetic poly(dc-da).(dg-dt)	
CC	probe. Over 100 repeat blocks were isolated. The inserts from the clones	
CC	were amplified by primers T65798-T66047. Those clones where the repeat	
CC	sequence has been determined are shown in T65704-797. This repeat	
CC	sequence is from the marker clone Mf669 which contains the repeat	
CC	sequence having the formula: (AC)18.5A(AC)3.	
SQ	Sequence 44 BP: 23 A; 21 C; 0 G; 0 U;	
Query Match 0.1%; Score 38; DB 1; Length 44;		
Best Local Similarity 100.0%; Pred. No. 0.00062;		
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
OY	1 CACACACACACACACACACACACACACAA 38	
DB	2 CACACACACACACACACACACACACACAA 39	
RESULT	2	
ID	T65773	
AC	T65773; standard; DNA; 51 BP.	
DT	17-JUN-1997 (first entry)	
DE	Repeat sequence from polymorphic marker clone Mf6102.	
KW	Polymorphism; repeat sequence; genetic marker; primer; amplification;	
KW	PCR; polymerase chain reaction; paternity; maternity; human; pedigree;	

Linkage analysis; genetic disease; animal; plant; breeding; locus;
 Hybridisation; chromosome; ds.
 Homo sapiens.
 US5582979-A.
 10-DEC-1996.
 PF 21-APR-1989; 341562.
 PR 21-APR-1989; US-341562.
 PR 05-SEP-1991; US-754351.
 PR 04-APR-1994; US-222177.
 PA (MARS-) MARSHFIELD CLINIC.
 PI Weber JL.
 DR WPI: 97-042299/04.
 PR Detection of polymorphic genetic markers of the form
 (dc-da)n(dc-dt)n - using novel nucleic acid mols. as primers
 PS Claim 1; Column 13-14; 186pp; English.
 CC The invention relates to the isolation of polymorphic repeat sequences
 having the sequence (dc-da)n.(dc-dt)n which can be used as genetic
 markers. Primers based on these sequences can be used to detect these
 repeats. Especially for use in e.g. paternity or maternity testing,
 human genetic analysis such as linkage analysis of genetic disease,
 commercial animal or plant breeding or pedigree analysis. Clones
 containing the repeat sequences were isolated by hybridisation of
 chromosome-specific phage libraries with a synthetic poly(dc-da) (dc-dt)
 probe. Over 100 repeat blocks were isolated. The inserts from the clones
 were amplified by primers 165798-166047. These clones where the repeat
 sequence has been determined are shown in 165704-1797. This repeat
 sequence is from the marker clone Kd1102 which contains the repeat
 sequence having the formula: (AC)19AA(AC)2A.
 CC Sequence 51 BP; 27 A; 24 C; 0 G; 0 U;
 SO

Query Match 0.1%; Score 38; DB 1; Length 51;
 Best Local Similarity 100.0%; Pred. No. 0.00061;
 Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CACACACACACACACACACACACACACACACACAA 38
 DB 4 CACACACACACACACACACACACACACACACACAA 41

RESULT 3
 ID 033930 standard; DNA: 60 BP.
 AC 033930:
 DR 02-FEB-1993 (first entry)
 DE Microsatellite sequence from clone TGA1337.
 KW PCR: selection; primers: OPTIPRIM; breeding; cattle; parentage;
 KM genetic mapping; traits; amplification; ss.
 OS Bos taurus.
 PN WO9213102-A.
 PD 06-AUG-1992.
 PF 15-JAN-1992; 000340.
 PR 15-JAN-1991; US-642342.
 PA (GENM-) GENMARK.
 PI Georges M, Massey JM.
 DR WPI: 92-284684/34.
 PR Polymorphic bovine DNA markers - used in genetic identification,
 PT gene mapping, and selective breeding
 PS Table 7; Page 301; 517pp; English.
 CC The sequence is that of a bovine microsatellite sequence obt'd. by
 screening a library of bovine MboI DNA fragments of between
 250 and 500 bp with an (AC)15 and a (TC)15 oligonucleotide probe.
 CC One out of 50 clones cross-hybridised. Assuming independent
 CC distribution of microsatellites and MboI sites, the frequency of
 CC (T6)n >9 microsatellites in the bovine genome is estimated at >100.
 CC 000 The sequence information for ca. 230 such bovine microsatellites
 CC is summarised in the specification and indexed herein (see below).
 CC The sequences upstream and downstream of the microsatellite sequence
 CC were used to generate the required PCR primers for in vitro
 CC amplification of the corresp. microsatellite (using the program
 CC OPTIPRIM). The microsatellites may be used to identify individuals,
 CC for parentage testing, and in the genetic mapping of economic trait
 CC loci, or genes involved in the determination of economically important

traits esp. in cattle, to allow selective breeding.
 CC See also 033501-34437.
 CC Sequence 60 BP; 2 A; 0 C; 27 G; 31 T;
 SO

Query Match 0.1%; Score 37; DB 1; Length 60;
 Best Local Similarity 100.0%; Pred. No. 0.0013;
 Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 ACACACACACACACACACACACACACACACACAA 38
 DB 41 ACACACACACACACACACACACACACACACACAA 5

RESULT 4
 ID 033849 standard; DNA: 70 BP.
 AC 033849:
 DR 02-FEB-1993 (first entry)
 DE Microsatellite sequence from clone TGA26.
 KW PCR: selection; primers: OPTIPRIM; breeding; cattle; parentage;
 KM genetic mapping; traits; amplification; ss.
 OS Bos taurus.
 PN WO9213102-A.
 PD 06-AUG-1992.
 PF 15-JAN-1992; 000340.
 PR 15-JAN-1991; US-642342.
 PA (GENM-) GENMARK.
 PI Georges M, Massey JM.
 DR WPI: 92-284684/34.
 PR Polymorphic bovine DNA markers - used in genetic identification,
 PT gene mapping, and selective breeding
 PS Table 7; Page 269; 517pp; English.
 CC The sequence is that of a bovine microsatellite sequence obt'd. by
 screening a library of bovine MboI DNA fragments of between
 250 and 500 bp with an (AC)15 and a (TC)15 oligonucleotide probe.
 CC One out of 50 clones cross-hybridised. Assuming independent
 CC distribution of microsatellites and MboI sites, the frequency of
 CC (T6)n >9 microsatellites in the bovine genome is estimated at >100.
 CC 000 The sequence information for ca. 230 such bovine microsatellites
 CC is summarised in the specification and indexed herein (see below).
 CC The sequences upstream and downstream of the microsatellite sequence
 CC were used to generate the required PCR primers for in vitro
 CC amplification of the corresp. microsatellite (using the program
 CC OPTIPRIM). The microsatellites may be used to identify individuals,
 CC for parentage testing, and in the genetic mapping of economic trait
 CC loci, or genes involved in the determination of economically important
 CC traits esp. in cattle, to allow selective breeding.
 CC See also 033501-34437.
 CC Sequence 70 BP; 1 A; 1 C; 33 G; 35 T;
 SO

Query Match 0.1%; Score 37; DB 1; Length 70;
 Best Local Similarity 100.0%; Pred. No. 0.0013;
 Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CACACACACACACACACACACACACACACACACAA 37
 DB 37 CACACACACACACACACACACACACACACACACAA 1

RESULT 5
 ID 033828 standard; DNA: 36 BP.
 AC 033828:
 DR 02-FEB-1993 (first entry)
 DE Microsatellite sequence from clone TGA23.
 KW PCR: selection; primers: OPTIPRIM; breeding; cattle; parentage;
 KM genetic mapping; traits; amplification; ss.
 OS Bos taurus.
 PN WO9213102-A.
 PD 06-AUG-1992.
 PF 15-JAN-1992; 000340.

PS Table 7: Page 189; 517bp. Bovine microsatellite sequence obt'd. by
CC The sequence is that of a bovine microsatellite sequence obtained by
CC screening a library of bovine MboI DNA fragments of between
CC 250 and 500 bp with an (AC)₁₅ and a (TC)₁₅ oligonucleotide probe.
CC One out of 50 clones cross-hybridised, assuming independent
CC distribution of microsatellites and MboI sites, the frequency of
CC (Tn)₉ microsatellites in the bovine genome is estimated at >100,
CC 000. The sequence information for ca. 230 such bovine microsatellites
CC is summarised in the specification and indexed herein (see below).
CC The sequences upstream and downstream of the microsatellite sequence
CC were used to generate the required PCR primers for *in vitro*
CC amplification of the corresp. microsatellite (using the program
CC OPRIMR). The microsatellites may be used to identify individuals,
CC for parentage testing, and in the genetic mapping of economic traits
CC local, or genes involved in the determinism of economically important
CC traits esp. in cattle, to allow selective breeding.
CC See also Q33501-34437.
SQ Sequence 39 BP; 0 A; 0 C; 19 G; 20 T;

	0.1%;	Score 36;	DB 1;	Length 39;
Query Match	Similarity	100.0%;	Pred. No. 0.003;	
Best Local	Matches	36;	Conservative	0; Mismatches
				0; Indels
				Gaps
OY	1	CACACACACACACACACACACACACACACA	36	
Db	38	CACACACACACACACACACACACACACACA	3	

RESULT 15
ID 033807/C
AC 033807 standard; DNA: 39 BP.
DT 033807;
DE 02-FEB-1993 (first entry)
KW Microsatellite sequence from clone TGLA213.
KW PCR; selection; Primers; OPR19M; breeding; cattle; parentage;
KW genetic mapping; traits; amplification; ss.
OS Bos taurus.
PN M09213102-A.
PD 06-AUG-1992.
PR 15-JAN-1992; U000340.
PR 15-JAN-1991; US-642342.
PR (GENM-) GENMARK.
PI Georges W, Massey JM;
DR WJ; 92-284684/34.
PT Polymorphic bovine DNA markers - used in genetic identification,
PT gene mapping, and selective breeding.
PS Table 7, Page 252; 51pp; English.
CC The sequence is that of a bovine microsatellite sequence obtd. by
CC screening a library of bovine MboI DNA fragments of between
CC 250 and 500 bp with an (AC)₁₅ and a (TC)₁₅ oligonucleotide probe.
CC One out of 50 clones cross-hybridised. Assuming independent
CC distribution of microsatellites and MboI sites, the frequency of
CC (76)n > microsatellites in the bovine genome is estimated at >100,
CC 000. The sequence information for ca. 230 such bovine microsatellites
CC is summarised in the specification and indexed herein (see below).
CC The sequences upstream and downstream of the microsatellite sequence
CC were used to generate the required PCR primers for in vitro
CC amplification of the corresp. microsatellite (using the program
CC OPTIRIV). The microsatellites may be used to identify individuals,
CC for parentage testing, and in the genetic mapping of economic trait
CC loci, or genes involved the determinism of economically important
CC traits esp. in cattle, to allow selective breeding.
CC See also Q33501-34437.
SQ Sequence 39 BP: 0 A; 0 C; 20 G; 19 T;

```
Query Match          0.18; Score 36; DB 1; Length 39;  
Best Local Similarity 100.0%; Pred.No. 0.003;  
Matches 36; Conservative 0; Mismatches 0; Indels 0  
  
      1 CACGACACACACACACACACACACACACACA 36  
      |||
```

Db 39 CACACACACACACACACACACACACACACA
Search completed: March 29, 2000, 13:08:06
Job time: 33484 sec

Mon Apr 3 08:24:16 2000

us-08-852-495c-1_copy_1_30000.rng

Page 7



9
1
2

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 29, 2000, 10:09:37 ; Search time 8959.53 Seconds

(without alignments)
-6778.337 Million cell updates/sec

Title: US-08-852-495c-1_COPY_25000_45000

Perfect score: 20001

Sequence: 1 AGTGCAGTACTGCGACCTC.....CAAGTTAGACATTTT 20001

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 821193 seqs, -1518192014 residues

Word size : 0

Total number of hits satisfying chosen parameters: 345848

Minimum DB seq length: 8

Maximum DB seq length: 105

Post-processing: Listing first 45 summaries

Database : GenDbml:*
1: gb_ba1:*
2: gb_ba2:*
3: gb_om:*
4: gb_ov:*
5: gb_pat:*
6: gb_ph:*
7: gb_pl1:*
8: gb_pl2:*
9: gb_pl3:*
10: gb_pr1:*
11: gb_pr2:*
12: gb_pr3:*
13: gb_ro:*
14: gb_sts:*
15: gb_sy:*
16: gb_un:*
17: gb_vl:*
18: em_fun:*
19: em_hum1:*
20: em_hum2:*
21: em_hum3:*
22: em_om:*
23: em_or:*
24: em_ov:*
25: em_pat:*
26: em_ph:*
27: em_pl:*
28: em_ro:*
29: em_sy:*
30: em_un:*
31: em_vl:*
32: gb_htg1:*
33: gb_htg2:*
34: gb_in1:*
35: gb_in2:*
36: em_ba1:*
37: em_ba2:*
38: em_hum3:*
39: em_hum4:*
40: gb_pr4:*
41: gb_htg3:*
42: gb_htg4:*
43: gb_htg5:*

44: gb_htg6:*
45: gb_htg7:*
46: em_htg1:*
47: em_htg2:*
48: em_htg3:*
49: em_hum5:*
50: gb_pl3:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	89	0.4	96	13 G31304	G31304 sy899g1-19
2	48	0.2	56	5 I31237	I31237 Sequence 14
3	46	0.2	85	10 HUMHIS1PR	M26162 Homo sapien
4	43	0.2	54	5 AR002288	AR002288 Sequence
5	43	0.2	54	5 AR053139	AR053139 Sequence
6	42	0.2	69	13 HUMUT578A	L39139 Human STS U
7	40	0.2	40	5 A68621	A68621 Sequence 1
8	40	0.2	101	10 S79560	S79560 HRX (partl
9	40	0.2	101	10 S79561	S79561 dhrx (partl
10	39	0.2	39	5 I31174	I31174 Sequence 86
11	39	0.2	39	5 I82155	I82155 Sequence 1
12	39	0.2	75	13 AF042304	AF042304 Homo sapl
13	39	0.2	102	39 SHPDINCRD	L01534 Sheep dlnuc
14	38	0.2	60	5 AR051487	AR051487 Sequence
15	38	0.2	66	5 AR002280	AR002280 Sequence
16	38	0.2	66	5 AR053141	AR053141 Sequence
17	38	0.2	76	5 AR051499	AR051499 Sequence
18	37	0.2	59	9 HSDXS30A3	X54497 H.sapiens 3
19	37	0.2	73	13 AF042303	AF042303 Homo sapl
20	36	0.2	40	5 I31240	I31240 Sequence 15
21	36	0.2	66	40 AF087511	AF087511 Homo sapl
22	35	0.2	35	5 A25212	A25212 Inter-Alu s
23	35	0.2	35	5 E09140	E09140 Synthetic D
24	35	0.2	45	5 I31245	I31245 Sequence 15
25	35	0.2	80	9 HUMBRKPAD	M36134 Human alpha
26	35	0.2	80	9 HUMBRKPAE	M36135 Human alpha
27	35	0.2	85	5 AR051522	AR051522 Sequence
28	34	0.2	40	5 A68621	A68621 Sequence 1
29	34	0.2	70	13 HUMUT5148A	L31067 Human STS U
30	34	0.2	87	12 MMYIMV27	X89144 M.musculus
31	33	0.2	35	5 A25212	A25212 Inter-Alu s
32	33	0.2	35	24 E09140	E09140 Synthetic D
33	33	0.2	55	9 HSDXS30A1	X54495 H.sapiens 3
34	33	0.2	70	12 MUSMABAD02	K02420 Mouse MHC c
35	33	0.2	80	5 HUMBRKPAE	M36133 Human alpha
36	33	0.2	73	5 A08915	A08915 H.sapiens (
37	32	0.2	83	13 AU025339	AU025339 Rattus no
38	32	0.2	90	9 HUMCDLREPL	K03555 Human low d
39	32	0.2	90	13 G37815	G37815 ACT1 Plasm
40	32	0.2	100	11 HS067848	U67848 Human beta-
41	32	0.2	102	5 A08911	A08911 H.sapiens (
42	32	0.2	104	13 AU046973	AU046973 Rattus no
43	31	0.2	60	11 HSCBFB5108	AF084962 Homo sapl
44	31	0.2	62	10 S52152S31	S52228 CD11b-1enke
45	31	0.2	71	13 HUMUT741A	L29638 Human STS U

ALIGNMENTS

RESULT 1
G31304
LOCUS
DEFINITION sy899g1-19 Human (A.Gnlrke) Homo sapiens STS genomic, sequence
tagged site.
ACCESSION G31304

Mon Apr 3 08:24:22 2000

us-08-852-495c-1_copy_25000_45000.rge

Page 2

VERSION	G31304.1	GI:1871333
KEYWORDS	STS.	
SOURCE	human.	
ORGANISM	Homo sapiens	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.	
AUTHORS	1 (bases 1 to 96) Lauer, P., Meyer, N.C., Prass, C.E., Searnes, S.M., Wolff, R.K. and Gnirke, A.	
TITLE	Clone-Contig and STS maps of the hereditary hemochromatosis region on human chromosome 6p21.3-6p22	
JOURNAL MEDLINE	Genome Res. 7 (5), 457-470 (1997)	
COMMENT	97294058	
	CDR: CDR:5581195	

STS size: 77
PCR profile:

Denaturation: 92 degrees C for 20 seconds
Annealing: 60 degrees C for 45 seconds
Polymerization: 72 degrees C for 60 seconds
PCR Cycles: 35
Thermal Cycler: MJ Research PTC-200

Template:	30-200 ng
Primer:	each 0.8 μ M
dNTPs:	each 200 μ M
Taq Polymerase:	0.05 units/ μ l
Total Vol:	12 μ l

buffer:	
MgCl2:	1.5 mM
KCl:	50 mM
Tris-HCl:	10 mM
pH:	8.3
gelatin:	0.0018 (w/v)
Location/Qualifiers	
source	1..96

```
STS /clone_lib="Human (A.G1nrke)"
primer_bind 9..85
primer_bind 9..30
primer_bind complement(70..85)
BASE COUNT 22 a 29 c 23 g 22 t
ORIGIN
```

Query Match 0.48; Score 89; DB 13; Length 96;
Best Local Similarity 100.0%; Pred. No. 1.9e-34;
Matches 89; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

0y 15871 CACCTGGTGTCCCAAGATATTAATGAGAAAAATGCTCCCATGATGCAAGATCCC 15930
Db 1 CCTCTGGTGTCCCAAGATATTAATGAGAAAAATGCTCCCATGATGCAAGATCCC 60
0y 15931 TCTGCCCCCTTCCCACTGATGCCCCCTGGGG 15959
Db 61 TCTGCCCCCTTCCCACTGATGCCCCCTGGGG 89

RESULT 2	
LOCUS	I31237 56 bp DNA
DEFINITION	Sequence 149 from patent US 5582979.
ACCESSION	I31237
VERSION	I31237.1 GI:1822028
06-FEB-1997	

KEYWORDS	Unknown.
SOURCE	Unknown.
ORGANISM	Unclassified.
REFERENCE	1 (bases 1 to 56)
AUTHORS	Weber, J.L.
TITLE	Length polymorphisms in (dc-da).sub.n.(dg-dt).sub.n sequences and method of using the same
JOURNAL	Patent: US 5582979-A 149 10-DEC-1996;
FEATURES	Location/Qualifiers
SOURCE	1..56
BASE COUNT	28 a 27 c 0 g 1 t
ORIGIN	/organism="unknown"

COMPR, 1565
32556 32859
OF T 1

```

Query Match          0.28; Score 48; DB 5; Length 56;
Best Local Similarity 100.0%; Pred. No. 3e-13;
Matches      48; Conservative    0; Mismatches     0; Indels   0; Gaps
              32512
QY 7513 GTGTGTCGTGTCGTGTCGTATGCTGGTGGTGTCGTGTCGTCT 7550
      |||.....|..|||.....|..|||.....|..|||.....|..
Db  56 GTGTGTCGTGTCGTGTCGTATGCTGGTGGTGTCGTGTCGTCT 9

```

LOCUS	HUMHIS1PR	85 bp	DNA	PRI	01-AUG-1995
DEFINITION	Homo sapiens histone 1 gene, promoter region.				
ACCESSION	U61652				
RESULT	3				

```

RECESSION  M26162.1  GI:514937
VERSION      histone 1; promoter
KEYWORDS     Homo sapiens DNA.
SOURCE       Homo sapiens

```

ORGANISM *Homo sapiens*
 Metazoa: Chordata; Vertebrata; Mammalia;
 Eukaryota; Metazoa; Craniata; Vertebrata;
 Eutheria; Primates; Catarrhini; Hominoidea;
 1 (bases 1 to 85)
 REFERENCE
 1. (bases 1 to 85)

REFERENCE	AUTHORS	TITLE	JOURNAL
1 (bases 1 to 85)	Gallinari, P., La Bella, F., and Heintz, N.	Characterization and purification of HIRF2, a novel CCAR-binding protein that interacts with a histone H1 subtype-specific consensus element	Mol. Cell. Biol. 9 (4), 1566-1575 (1989)
2			Mol. Cell. Biol. 9 (4), 1566-1575 (1989)

REELLINE	09201/04
COMMENT	On Jul 23, 1994 this sequence version replaced g1:341438
FEATURES	Location/Qualifiers
SOURCE	1. .85

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_line="HeLa"
28..39
protein_bind
bound_motety="SP1"
53..70
CAAT_signal

```

gene	/Gene="HIS1"
53. .85	/Gene="HIS1"
77. .85	

BASE COUNT	26 a	27 c	18 g	14 t
ORIGIN				

```

Query Match      0.28; Score 46; DB 10; Length 85;
Best Local Similarity 100.0%; Pred. No. 3.2e-12;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

0y 1025 GTGCGCTAGTGATGCACCAATCACAGGGCCCTACCCATATATAG 10
 |||||
 Db 40 GTCCCTGATGATGCACCAATCACAGGGCCCTACCCATATATAG 85

RESULT	4			
AR002288/c				
LOCUS	AR002288	54 bp	DNA	PAT
				04-DEC-1998

[illegible]

```

TITLE
JOURNAL
COMMENT
Melis, R., Lawrence, E., Moore, M., Holik, P. R., Carlson, M., Zhao, X.,
Robertson, M., Bradley, P., Elsner, T., Tingey, A., Lalouel, J.-M. and
White, R.
Genetic and physical mapping of simple sequence repeat containing
sequence tagged sites from the human genome
Unpublished (1994)
Submitted by: Utah Center for Human Genome Research University of
Utah, Dept. of Human Genetics
2160 Eccles Institute of Human Genetics
Salt Lake City, UT 84112
e-mail: sts@corona.med.utah.edu
Primer A: AGTCGAGACGACGCTGGC
Primer B: AGGTGGCAGAAATGCGCATC
End to Label: Primer A
PCR Profile:
Initial Denaturation: 94C 300sec
Cycles Denaturation Annealing Extension 5 94
C 10 sec. 62 C 10 sec. 72 C 20 sec. 30
58 C 10 sec. 72 C 20 sec. Mg++: 1.50 mM
Gel: Acrylamide 7%, Formamide 32%, Urea 34%
Alleles: 1.
FEATURES
source
1..69
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="19"
14..32
/note="STS UM578 5' end"
/evidence=experimental
BASE COUNT 26 a 16 c 14 g 13 t
ORIGIN
Query Match 0.2%; Score 42; DB 13; Length 69;
Best Local Similarity 100.0%; Pred. No. 3.7e-10;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 10109 TTGAGTCAGAGAGTTCGAGACCGCTGGCCAACTGTGTGA 10150
|||||
|||||
Db 3 TTGAGTCAGAGAGTTCGAGACCGCTGGCCAACTGTGTGA 44
|||||
|||||
RESULT 7
A68621/c 40 bp DNA PAT 06-MAY-1999
LOCUS A68621
DEFINITION Sequence 1 from Patent WO9801573.
ACCESSION A68621
VERSION A68621.1 GI:4759648
KEYWORDS
SOURCE
ORGANISM
unidentified.
unidentified.
unclassified.
1 (bases 1 to 40)
REFERENCE
Renwick, M. A., Lariouev, V. I., Koudrina, N. Y. and Perkins, E. L.
TRANSFORMATION-ASSOCIATED RECOMBINATION CLONING
Patent: WO 9801573-A 15-JAN-1998;
US HEALTH (US)
FEATURES
source
1..40
/location/Qualifiers
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 7 a 12 c 13 g 8 t
ORIGIN
Query Match 0.2%; Score 40; DB 5; Length 40;
Best Local Similarity 100.0%; Pred. No. 4.1e-09;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 8640 TCGGCTCCCAAGTGTGGATTCAGAGCGTGAGCCACC 8679
|||||
|||||
Db 40 TCGGCTCCCAAGTGTGGATTCAGAGCGTGAGCCACC 1

```

[illegible][illegible]

[illegible]

```

      .
      repeat_region
      BASE COUNT          40 a          42 c          9 g          11 t
      ORIGIN
Query Match              0.2%; Score 39; DB 3; Length 102;
Best Local Similarity   100.0%; Pred. No. 1.3e-08;
Matches                 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0
QY    7517 GTGTGCTGTGTGTGTGTGTATGTGTGTGTGTGTG 7555
      |||||||
      Db    69 GTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 31

RESULT 14
AR051487/c            AR051487        60 bp       DNA           PAT           29-SEP-1999
LOCUS                Sequence 57 from patent US 5830670.
DEFINITION            AR051487
ACCESSION             AR051487
VERSION               AR051487.1 GI:5974851
KEYWORDS
SOURCE                Unknown.
ORGANISM              Unclassified.
REFERENCE              1 (bases 1 to 60)
AUTHORS               de la Monte,S., and Wands,J.R.
TITLE                 Neural thread protein gene expression and detection of Alzheimer's
                     disease
JOURNAL               Patent: US 5830670-A 57 03-NOV-1998;
FEATURES
source               Location/Qualifiers
                     1..60
BASE COUNT          12 a          14 c          15 g          19 t
ORIGIN
Query Match              0.2%; Score 38; DB 5; Length 60;
Best Local Similarity   100.0%; Pred. No. 4.4e-08;
Matches                 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0
QY    12897 GGTTGAACCCTGTCCTACTAAAAATTACAAAATTAGC 12934
      |||||||
      Db    48 GGTTGAACCCTGTCCTACTAAAAATTACAAAATTAGC 11

RESULT 15
AR002290            AR002290        66 bp       DNA           PAT           04-DEC-1998
LOCUS                Sequence 29 from patent US 5741645.
DEFINITION            AR002290
ACCESSION             AR002290
VERSION               AR002290.1 GI:3963844
KEYWORDS
SOURCE                Unknown.
ORGANISM              Unclassified.
REFERENCE              1 (bases 1 to 66)
AUTHORS               Orr,H.T., Rannum,L.P.W., Chung,M., and Zoghbi,H.Y.
TITLE                 Gene sequence for spinocerebellar ataxia type I and method for
                     diagnosis
JOURNAL               Patent: US 5741645-A 29 21-APR-1998;
FEATURES
source               Location/Qualifiers
                     1..66
BASE COUNT          12 a          0 c          31 g          23 t
ORIGIN
Query Match              0.2%; Score 38; DB 5; Length 66;
Best Local Similarity   100.0%; Pred. No. 4.3e-08;
Matches                 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Mon Apr 3 08:24:22 2000

us-08-852-495c-1_copy_25000_45000.rge

Page 6

```
OY      7513 GTCGTGCTGTCGTGTCGTGTCGTATGTCGTGTCGT    7550
          |||||
DB      29 GTCGTGCTGTCGTGTCGTGTCGTATGTCGTGTCGT    66
```

Search completed: March 29, 2000, 17:18:26
Job time: 50055 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 29, 2000, 17:18:26 ; Search time 13410.2 Seconds
(without alignments)
-6792.946 Million cell updates/sec

Title: US-08-852-495c-1_COPY_40000_70000
Perfect score: 30001
Sequence: 1 GGTGACACGTCAGTCA.....AACACAGTTCAGAAATGA 30001

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 821193 seqs, -1518192014 residues

Word size : 0

Total number of hits satisfying chosen parameters: 345848

Minimum DB seq length: 8
Maximum DB seq length: 105

Post-processing: Listing first 45 summaries

Database : GenEmbl.*
1: gb_da1:*
2: gb_da2:*
3: gb_da3:*
4: gb_da4:*
5: gb_da5:*
6: gb_da6:*
7: gb_da7:*
8: gb_da8:*
9: gb_da9:*
10: gb_da10:*
11: gb_da11:*
12: gb_da12:*
13: gb_da13:*
14: gb_da14:*
15: gb_da15:*
16: gb_da16:*
17: gb_da17:*
18: gb_da18:*
19: gb_da19:*
20: gb_da20:*
21: gb_da21:*
22: gb_da22:*
23: gb_da23:*
24: gb_da24:*
25: gb_da25:*
26: gb_da26:*
27: gb_da27:*
28: gb_da28:*
29: gb_da29:*
30: gb_da30:*
31: gb_da31:*
32: gb_da32:*
33: gb_da33:*
34: gb_da34:*
35: gb_da35:*
36: gb_da36:*
37: gb_da37:*
38: gb_da38:*
39: gb_da39:*
40: gb_da40:*
41: gb_da41:*
42: gb_da42:*
43: gb_da43:*

44: gb_da44:*
45: gb_da45:*
46: gb_da46:*
47: gb_da47:*
48: gb_da48:*
49: gb_da49:*
50: gb_da50:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	89	0.3	96	13 G31304	G31304 sy899g1-19
2	46	0.2	92	8 AF062763	AF062763 Glycine m
3	45	0.1	99	13 G37948	G37948 C4W78 plasm
4	44	0.1	99	13 G37948	G37948 C4W78 plasm
5	43	0.1	90	9 HUMDLRMA	M15365 Human low d
6	42	0.1	67	9 HUMALUDANCD	L36843 Homo sapien
7	41	0.1	90	9 HUMDLRMA	M15365 Human low d
8	41	0.1	92	8 AF062763	AF062763 Glycine m
9	41	0.1	104	9 HUMALCE272	M87899 Human carc1
10	39	0.1	39	5 I82155	I82155 Sequence 1
11	39	0.1	60	5 I88893	I88893 Sequence 11
12	39	0.1	80	9 HUMBRKFFAA	M36131 Human alpha
13	38	0.1	40	5 A68622	A68622 Sequence 2
14	38	0.1	58	5 AR063068	AR063068 Sequence
15	38	0.1	62	10 S52152531	S52228 CD11b-Leuko
16	38	0.1	70	5 A49117	A49117 Sequence 12
17	38	0.1	75	5 A49118	A49118 Sequence 13
18	38	0.1	80	5 A18777	A18777 5' terminus
19	38	0.1	80	5 AR001312	AR001312 Sequence
20	38	0.1	90	9 HUMDLRFL	K03555 Human low d
21	37	0.1	37	5 I29931	I29931 Sequence 44
22	37	0.1	40	5 A48799	A48799 Sequence 6
23	37	0.1	50	5 I36502	I36502 Sequence 1
24	37	0.1	55	5 I29928	I29928 Sequence 41
25	37	0.1	69	5 I89934	I89934 Sequence 7
26	37	0.1	69	5 I89944	I89944 Sequence 22
27	37	0.1	72	10 S76508	S76508 DIS8 (B) (h
28	37	0.1	72	10 S76510	S76510 DIS8 (B) (h
29	37	0.1	73	5 A08915	A08915 H. sapiens (
30	37	0.1	77	5 A08908	A08908 H. sapiens (
31	37	0.1	80	5 A08907	A08907 H. sapiens (
32	37	0.1	80	9 HUMBRKFFAB	M36132 Human alpha
33	37	0.1	80	9 HUMBRKFFAC	M36133 Human alpha
34	37	0.1	80	9 HUMBRKFFAC	M36133 Human alpha
35	37	0.1	84	5 AR038854	AR038854 Sequence
36	37	0.1	84	5 AR050387	AR050387 Sequence
37	37	0.1	90	5 E12579	E12579 Probe. 6/19
38	37	0.1	90	5 E12580	E12580 Probe. 6/19
39	37	0.1	90	5 E13639	E13639 Probe. 6/19
40	37	0.1	90	5 E17120	E17120 DNA probe.
41	37	0.1	91	34 DDICACTN8A	M25215 Dictyostel1
42	37	0.1	96	5 A08909	A08909 H. sapiens (
43	37	0.1	98	5 I49625	I49625 Sequence 42
44	37	0.1	102	5 A08911	A08911 H. sapiens (
45	37	0.1	104	5 A08910	A08910 H. sapiens (

ALIGNMENTS

RESULT 1
LOCUS G31304 96 bp DNA
DEFINITION sy899g1-19 Human (A.Gnlrke) Homo sapiens STS genomic, sequence
tagged site.
ACCESSION G31304

Query Match 0.1%; Score 43; DB 9; Length 90;
 Best Local Similarity 100.0%; Pred. No. 5.6e-11;
 Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9929 CCTGAGCTCCCAAGTGTGGATTACAGCATGACCCCA 9971
 |||||||
 DB 38 CCTGAGCTCCCAAGTGTGGATTACAGCATGACCCCA 80

RESULT 6
 HUMALUNCD 67 bp DNA PRI 08-OCT-1994
 LOCUS Homo sapiens 4000 year old remains from Nekht-ankh Alu repeat
 DEFINITION fragment 12:5.
 ACCESSION L36843.1 GI:556196
 VERSION L36843
 KEYWORDS Alu repeat.
 SOURCE Homo sapiens (individual isolate 4000 year old remains from Nekht-ankh) liver DNA.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (sites)
 AUTHORS Pabo, S.
 TITLE Ancient DNA: extraction, characterization, molecular cloning, and enzymatic amplification
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 86 (6), 1939-1943 (1989)
 MEDLINE 89184542
 FEATURES location/Qualifiers
 source 1..67
 /organism="Homo sapiens"
 /isolate="4000 year old remains from Nekht-ankh"
 /db_xref="taxon:9606"
 /tissue_type="liver"
 /rpt_family="Alu"
 repeat_region 15 a 18 c 23 g 11 t

Query Match 0.1%; Score 42; DB 9; Length 67;
 Best Local Similarity 100.0%; Pred. No. 1.9e-10;
 Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20914 GTGGCTCACACCTGTATCCAGCATTGGAGGCTGAGG 20955
 |||||||
 DB 9 GTGGCTCACACCTGTATCCAGCATTGGAGGCTGAGG 50

RESULT 7
 HUMDLRM/c 90 bp DNA PRI 11-JAN-1995
 LOCUS Human low density lipoprotein receptor mutant gene recombination site.
 DEFINITION
 ACCESSION M15365.1 GI:187107
 VERSION M15365.1
 KEYWORDS Alu repeat; LDL receptor; cell surface protein; crossover; recombination.
 SOURCE Human (FH 295) fibroblast DNA, clone p295.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 90)
 AUTHORS Lehman, M.A., Goldstein, J.L., Russell, D.W. and Brown, M.S.
 TITLE Duplication of seven exons in LDL receptor gene caused by Alu-Alu recombination in a subject with familial hypercholesterolemia
 JOURNAL Cell 48 (5), 827-835 (1987)
 MEDLINE 87131094
 COMMENT Clean copy of sequence kindly provided by M. Lehman (22-APR-1987). Individual FH 295 carries two mutant LDL receptor alleles. The allele below was inherited from the father and includes a duplication of exons 2 through 8 that was presumably created by

unequal chromosomal crossover involving Alu repeats in introns 1 and 8. The exact site of recombination cannot be determined because the sequences of the Alu repeats in introns 1 and 8 of normal alleles are identical over a span of 26 nucleotides at the recombination site

FEATURES
 source 1..90
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /map="19p13.3"
 1..43
 /gene="LDLR"
 /note="LDLR"
 /gene="LDLR"
 /note="LDLR intron 8; G00-119-362"
 43..68
 /organism="Homo sapiens"
 68..>90
 /note="LDLR duplicated intron 1 (no splice consensus at 68); putative; does not fit consensus"
 18 a 33 c 19 g 20 t
 BASE COUNT 18 a 33 c 19 g 20 t
 ORIGIN 1 bp upstream of NheI site; chromosome 19p13.2-p13.1.

Query Match 0.1%; Score 41; DB 9; Length 90;
 Best Local Similarity 100.0%; Pred. No. 6.3e-10;
 Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 23593 GTGGCTCAGCCCTGTATCCAGCATTGGAGGCTGAGG 23633
 |||||||
 DB 78 GTGGCTCAGCCCTGTATCCAGCATTGGAGGCTGAGG 38

RESULT 8
 AF062763/c 92 bp DNA PLN 02-AUG-1998
 LOCUS Glycine max microsatellite AC21 repeat region.
 DEFINITION
 ACCESSION AF062763
 VERSION AF062763.1 GI:3372757
 KEYWORDS soybean.
 SOURCE Glycine max
 ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Glycine.
 REFERENCE 1 (bases 1 to 92)
 AUTHORS Peakall, R., Gilmore, S., Keys, W., Morgante, M. and Rafalski, A.
 TITLE Cross-species amplification of soybean (Glycine max) simple-sequence-repeats (SSRs) within the genus and other legume genera: implications for the transferability of SSRs in plants
 JOURNAL Mol. Biol. Evol. (1998) in press
 2 (bases 1 to 92)
 REFERENCE Peakall, R., Gilmore, S., Keys, W., Morgante, M. and Rafalski, A.
 TITLE Direct Submission
 JOURNAL Submitted (04-MAY-1998) Division of Botany and Zoology, Australian National University, Canberra, ACT 0200, Australia
 FEATURES location/Qualifiers
 source 1..92
 /organism="Glycine max"
 /db_xref="taxon:3847"
 /note="microsatellite AC21"
 repeat_region 6..57
 /rpt_type="tandem"
 /rpt_unit="ta
 BASE COUNT 48 a 5 c 4 g 35 t
 ORIGIN

Query Match 0.1%; Score 41; DB 8; Length 92;
 Best Local Similarity 100.0%; Pred. No. 6.3e-10;
 Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
BASE COUNT
ORIGIN

unidentified.
unidentified.
unclassified.
1 (bases 1 to 40)
Resnick, M.A., Laktionov, V.L., Koupina, N.Y. and Perkins, E.L.
TRANSFORMATION-ASSOCIATED RECOMBINATION CLONING
Patent: WO 9801573-A 15-JAN-1998;
US HEALTH (US)
Location/Qualifiers
1..40
/organism="unidentified"
/db_xref="taxon:32644"
9 a 8 c 19 g 4 t

Query Match 0.1%; Score 38; DB 5; Length 40;
Best Local Similarity 100.0%; Pred. No. 2.4e-08;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 8613 GAGCTGAGCAGAGACATCGTGAACCGGAGCG 8650
Db 2 GAGCTGAGCAGAGACATCGTGAACCGGAGCG 39

RESULT 14
LOCUS AR063068 58 bp DNA PAT 29-SEP-1999
DEFINITION Sequence 42 from patent US 5844075.
ACCESSION AR063068
VERSION AR063068.1 GI:5990759
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
AUTHORS 1 (bases 1 to 58)
Kawakami, Y. and Rosenberg, S.A.
TITLE Melanoma antigens and their use in diagnostic and therapeutic methods
JOURNAL Patent: US 5844075-A 42 01-DEC-1998;
FEATURES Location/Qualifiers
1..58
/organism="unknown"
BASE COUNT 3 a 6 c 9 g 40 t
ORIGIN

Query Match 0.1%; Score 38; DB 5; Length 58;
Best Local Similarity 100.0%; Pred. No. 2.4e-08;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 13600 CTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 13637
Db 18 CTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 55

RESULT 15
LOCUS S52152531 62 bp DNA PRI 08-MAY-1993
DEFINITION CD11b-leukocyte integrin alpha chain [human, Genomic, 62 nt.
ACCESSION S52228
VERSION S52228.1 GI:263048
KEYWORDS
SEGMENT 31 of 31
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 62)
Fleming, J.C., Pahl, H.L., Gonzalez, D.A., Smith, T.F. and Tenen, D.G.
TITLE Structural analysis of the CD11b gene and phylogenetic analysis of

the alpha-integrin gene family demonstrate remarkable conservation of genomic organization and suggest early diversification during evolution
JOURNAL U. Immunol. 150 (2), 480-490 (1993)
MEDLINE 93123748
REMARK Genbank staff at the National Library of Medicine created this entry [NCBI g1dbsg 121959] from the original journal article.
Map location: 16.
Location/Qualifiers
1..62
/organism="Homo sapiens"
/db_xref="taxon:9606"
51 a 2 c 1 g 8 t

BASE COUNT 51 a 2 c 1 g 8 t
ORIGIN

Query Match 0.1%; Score 38; DB 10; Length 62;
Best Local Similarity 100.0%; Pred. No. 2.4e-08;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 13601 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 13638
Db 58 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 21

Search completed: March 30, 2000, 00:22:34
Job time: 73503 sec

Mon Apr 3 08:24:25 2000

us-08-852-495c-1_copy_40000_70000.rge



2

7

•

1 CLASSIFICATION: 435
 2 PRIOR APPLICATION DATA:
 3 APPLICATION NUMBER: US 07/341,562
 4 FILING DATE: 21-APR-1989
 5 ATTORNEY/AGENT INFORMATION:
 6 NAME: Sate, Charles S.
 7 REGISTRATION NUMBER: 30,492
 8 REFERENCE/DOCKET NUMBER: 09865.601
 9 TELECOMMUNICATION INFORMATION:
 10 TELEPHONE: (608) 831-2100
 11 TELEFAX: (608) 831-2106
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 51

	Query Match	0.1%	Score 36;	DB 1;	Length 38;
	Best Local Similarity	100.0%	Pred. No.	0.00084;	
Matches	36;	Conservative	0;	Mismatches	0;
				Indels	0;
Gaps					0;
OY	1 CACACACACACACACACACACACACACACA	36			
Db	1 CACACACACACACACACACACACACACACA	36			

```

1      RESULT      5
2      US-08-222-177A-397
3      Sequence 397 Application US/08222177A
4      Patent No. 5582579
5      GENERAL INFORMATION:
6      APPLICANT: Weber, James L.
7      TITLE OF INVENTION: LENGTH POLYMORPHISMS IN
8      TITLE OF INVENTION: (dC-dA)n..(dG-dT)n SEQUENCES AND METHODS OF USING SAME
9      NUMBER OF SEQUENCES: 460
10     CORRESPONDENCE ADDRESS:
11     ADDRESSEE: Dewitt Ross & Stevens, S.C.
12     STREET: 8000 Excelsior Drive, Suite 401
13     CITY: Madison
14     STATE: Wisconsin
15     COUNTRY: USA
16     ZIP: 53717-1914
17     COMPUTER READABLE FORM:
18     MEDIUM TYPE: Floppy disk
19     COMPUTER: IBM PC compatible
20     OPERATING SYSTEM: PC-DOS/MS-DOS
21     SOFTWARE: PatentIn Release #1.0, Version #1.25
22     CURRENT APPLICATION DATA:
23     APPLICATION NUMBER: US/08/222,177A
24     FILING DATE:
25     CLASSIFICATION: 435
26     PRIOR APPLICATION DATA:
27     APPLICATION NUMBER: US 07/341,562
28     FILING DATE: 21-APR-1989
29     ATTORNEY/AGENT INFORMATION:
30     NAME: Sara, Charles S.
31     REGISTRATION NUMBER: 30,492
32     REFERENCE/DOCKET NUMBER: 09865, 601
33     TELECOMMUNICATION INFORMATION:
34     TELEPHONE: (608) 831-2100
35     TELEFAX: (608) 831-2106
36     TELEX:
37     INFORMATION FOR SEQ ID NO: 397:
38     SEQUENCE CHARACTERISTICS:
39     LENGTH: 38 base pairs
40     TYPE: nucleic acid

```

```

; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-222-177A-397

```

```

Query Match      0.1%; Score 36; DB 1; Length 38;
Best Local Similarity 100.0%; Pred. No. 0 00084;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy 1 CACACACACACACACACACACACACACACA 36
 |||||
Dd 2 CACACACACACACACACACACACACACACA 37

RESULT 6
US-08-222-177A-137
Sequence 137, Application US/08222177A
Patent No. 7,800,070

```

: APPLICANT: Weber, James L.
: TITLE OF INVENTION: LENGTH POLYMORPHISMS IN
: TITLE OF INVENTION: (dc-da)n.(dg-dt)n SEQUENCES AND METHODS OF USING SAME
: TITLE OF INVENTION:
:

```

```

CORRESPONDENCE ADDRESS:
ADDRESS: Dewitt Ross & Stevens, S.C.
STREET: 8000 Excelsior Drive, Suite 401
CITY: Charleston

```

COUNTRY: USA
ZIP: 53717-1914
COMPUTER READABLE FORM:
MEDICAL TYPE: 510000

```

:
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
:
: CURRENT APPLICATION DATA:
:
: ADDITIONAL NUMBER: HC/09/000 177A

```

```

; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/341,562
; FILING DATE: 31-AND-1080

```

NAME: Sara, Charles S.
REGISTRATION NUMBER: 30,492
REFERENCE/DOCKET NUMBER: 09865.601

TELEFAX: (608) 831-2106
TELEX:
INFORMATION FOR SEQ ID NO: 137:

```

; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear

```

;
; IMMEDIATE SOURCE:
; CLONE: mfd29rs
US-08-222-177A-137

Query Match	0.1%;	Score 36;	DB 1;	Length 39;
Best Local Similarity	100.0%;	Pred. No. 0	00084;	
Matches 36;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

```
QY      1 CACACAGCACACACACACACACACACA   36  
        |||||  
Db       2 CACACAGCACACACACACACACACA   37
```

RESULT 7

Mon Apr 3 08:24:17 2000

us-08-852-495c-1_copy_1_30000.rni

Page 4

```

US-08-222-177A-119
; Sequence 119, Application US/08222177A
; Patent No. 5582979
GENERAL INFORMATION:
APPLICANT: Weber, James L.
TITLE OF INVENTION: LENGTH POLYMORPHISMS IN
NUMBER OF SEQUENCES: 460
CORRESPONDENCE ADDRESS:
ADDRESSEE: Demilt Ross & Stevens, S.C.
STREET: 8000 Excelsior Drive, Suite 401
City: Madison
STATE: Wisconsin
ZIP: 53717-1914
COMPUTER READABLE FORM:
COMPUTER TYPE: Floppy disk
MEDIUM TYPE: FLOPPY DISK
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/222,177A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/341,562
FILING DATE: 21-APR-1989
ATTORNEY/AGENT INFORMATION:
NAME: Sara, Charles S.
REGISTRATION NUMBER: 30,492
REFERENCE/DOCKET NUMBER: 09865, 601
TELECOMMUNICATION INFORMATION:
TELEPHONE: (608) 831-2100
TELEFAX: (608) 831-2106
TELEX:
INFORMATION FOR SEQ ID NO: 119:
SEQUENCE CHARACTERISTICS:
LENGTH: 40 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
IMMEDIATE SOURCE:
CLONE: mfd37s
US-08-222-177A-119
Query Match 0.18; Score 36; DB 1; Length 40;
Best Local Similarity 100.0%; Fred. No. 0.00084;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1 CACACACACACACACACACACACACACACACACA 36
| | | | | | | | | | | | | | | | | | | | | |
Db 2 CACACACACACACACACACACACACACACACACA 37
RESULT 8
US-08-222-177A-400
; Sequence 400, Application US/08222177A
; Patent No. 5582979
GENERAL INFORMATION:
APPLICANT: Weber, James L.
TITLE OF INVENTION: LENGTH POLYMORPHISMS IN
NUMBER OF SEQUENCES: 460
CORRESPONDENCE ADDRESS:
ADDRESSEE: Demilt Ross & Stevens, S.C.
STREET: 8000 Excelsior Drive, Suite 401
City: Madison
STATE: Wisconsin
ZIP: 53717-1914
COMPUTER READABLE FORM:

```

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/222,177A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/341,562
FILING DATE: 21-APR-1989
ATTORNEY/AGENT INFORMATION:
NAME: Sara, Charles S.
REGISTRATION NUMBER: 30,492
REFERENCE/DOCKET NUMBER: 09865,601
TELECOMMUNICATION INFORMATION:
TELEPHONE: (608) 831-2100
TELEFAX: (608) 831-2106

US-08-222-177A-400

MOLECULE TYPE: DNA (genomic)

STRANDEDNESS: double
TOPOLOGY: linear

INFORMATION FOR SEQ ID NO.:
SEQUENCE CHARACTERISTICS:
LENGTH: 40 base pairs
TYPE: nucleic acid
STRANDEDNESS: double

Query Match 0.1%; Score 36; DB 1; Length 40;
Best Local Similarity 100.0%; Pred. No. 0.00084;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CACACACACACACACACACACACACACACACACA 36
Db 2 CACACACACACACACACACACACACACACACACA 37

RESULT 9
US-08-222-177A-403
Sequence 403. Application US/08222177A
Patent No. 5582879
GENERAL INFORMATION:
APPLICANT: Weber, James L.
TITLE OF INVENTION: LENGTH POLYMORPHISMS IN
TITLE OF INVENTION: (dc-dn)n.(dg-dt)n SEQUENCES AND METHODS OF USING SAME
NUMBER OF SEQUENCES: 460
CORRESPONDENCE ADDRESS:
ADDRESSEE: Demilt, Ross & Stevens, S.C.
STREET: 8000 Excelsior Drive, Suite 401
CITY: Madison
STATE: Wisconsin
COUNTRY: USA
ZIP: 53717-1914
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/222,177A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/341,562
FILING DATE: 21-APR-1989
ATTORNEY/AGENT INFORMATION:
NAME: Sara, Charles S.
REGISTRATION NUMBER: 30,492
REFERENCE/DOCKET NUMBER: 09865,601
TELECOMMUNICATION INFORMATION:
TELEPHONE: (608) 831-2100
TELEFAX: (608) 831-2106

Mon Apr 3 08:24:17 2000

us-08-852-495c-1_copy_1_30000.mri

Page 8

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 29, 2000, 03:15:46 ; Search time 8180.08 seconds
(without alignments)
13847.043 Million cell updates/sec

Title: US-08-852-495c-1_COPY_1_30000
Perfect score: 30000
Sequence: 1 CACACACACACACACACACA.....AGCCATACAGAGTGGCCCC 30000

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 4538634 seqs, 1887831982 residues

Word size : 0

Total number of hits satisfying chosen parameters: 130662

Minimum DB seq length: 8
Maximum DB seq length: 105

Post-processing: Listing first 45 summaries

Database : EST:*

1:	em_est1:*
2:	em_est2:*
3:	em_est3:*
4:	em_est4:*
5:	em_est5:*
6:	em_est6:*
7:	em_est7:*
8:	em_est8:*
9:	em_est9:*
10:	em_est10:*
11:	em_est11:*
12:	em_est12:*
13:	em_est13:*
14:	em_est14:*
15:	em_est15:*
16:	em_est16:*
17:	em_est17:*
18:	em_est18:*
19:	em_est19:*
20:	gb_est1:*
21:	gb_est2:*
22:	gb_est3:*
23:	gb_est4:*
24:	gb_est5:*
25:	gb_est6:*
26:	gb_est7:*
27:	gb_est8:*
28:	gb_est9:*
29:	gb_est10:*
30:	gb_est11:*
31:	gb_est12:*
32:	gb_est13:*
33:	gb_est14:*
34:	gb_est15:*
35:	gb_est16:*
36:	gb_est17:*
37:	gb_est18:*
38:	gb_est19:*
39:	gb_est20:*
40:	gb_est21:*
41:	gb_est22:*
42:	gb_est23:*
43:	gb_est24:*

44:	gb_est25:*
45:	gb_est26:*
46:	gb_est27:*
47:	gb_est28:*
48:	gb_est29:*
49:	gb_est30:*
50:	gb_est31:*
51:	gb_est32:*
52:	em_est20:*
53:	em_est21:*
54:	em_est22:*
55:	em_est23:*
56:	em_est24:*
57:	em_est25:*
58:	em_est26:*
59:	gb_est33:*
60:	gb_est34:*
61:	gb_est35:*
62:	gb_est36:*
63:	gb_est37:*
64:	gb_est38:*
65:	em_est27:*
66:	em_est28:*
67:	em_est29:*
68:	em_est30:*
69:	gb_est39:*
70:	gb_est40:*
71:	gb_est41:*
72:	gb_est42:*
73:	gb_est43:*
74:	gb_est44:*
75:	em_est31:*
76:	em_est32:*
77:	em_est33:*
78:	em_est34:*
79:	gb_gss1:*
80:	gb_gss2:*
81:	gb_gss3:*
82:	gb_gss4:*
83:	em_gss1:*
84:	em_gss2:*
85:	em_gss3:*
86:	em_gss4:*
87:	gb_gss5:*
88:	gb_gss6:*
89:	gb_gss7:*
90:	gb_gss8:*
91:	gb_gss9:*
92:	em_gss5:*
93:	em_gss6:*
94:	em_gss7:*
95:	em_gss8:*
96:	em_gss9:*
97:	em_gss10:*
98:	em_gss11:*
99:	gb_gss10:*
100:	gb_gss11:*
101:	em_gss12:*
102:	gb_gss12:*
103:	gb_gss13:*
104:	gb_gss14:*
105:	gb_gss15:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
c 1	80	0.3	103	21	T77382
T77382 ydt2h12.r1					

2	73	0.2	95	21	R08388	AA500098	AA500098.y190a02.r
3	39	0.1	89	34	AA350598	T522775	AA500598.ya79d08.r
c	4	0.1	95	20	T522775	AA500598	T522775.ya79d08.r
c	5	0.1	101	90	AA006534	AA006534	AA006534.CIT-HSP-2
c	6	0.1	54	40	AA935225	AA935225	AA935225.co06b11.s
c	7	0.1	58	68	HSM010004	HSM010004	AA045154.Homo.sapi
c	8	0.1	89	50	FT4480	FT4480	HSPD10834.H
c	9	0.1	89	36	AA598741	AA598741	AA598741.ae49g12.s
c	10	0.1	102	81	BC62983	BC62983	CIT-HSP-669
c	11	0.1	58	81	B033985	B033985	CIT-HSP-21b5-u
c	12	0.1	80	44	AT262095	AT262095	gqr28c05.y
c	13	0.1	84	50	AT34654	AT34654	HSPE29754.H
c	14	0.1	89	28	AA078242	AA078242	AT078242.7H1605.C
c	15	0.1	89	36	AA596741	AA596741	ae49g12.s
c	16	0.1	97	28	AA077989	AA077989	7H12B06.C
c	17	0.1	99	29	AA144458	AA144458	mr080h5.r
c	18	0.1	101	63	AA1964456	AA1964456	4b6010G12
c	19	0.1	101	81	BL1560	BL1560	345315.TV.C
c	20	0.1	102	81	B47960	B47960	RPC111-3M4
c	21	0.1	103	104	AQ582106	AQ582106	RPC111-3M4
c	22	0.1	104	38	AA7963322	AA7963322	vs94e03.r
c	23	0.1	77	35	AA560791	AA560791	v133b06.r
c	24	0.1	85	81	B02832	B02832	CSRL-160e3-r
c	25	0.1	86	38	AA745546	AA745546	nv56g01.s
c	26	0.1	86	38	AA745546	AA745546	nv56g01.s
c	27	0.1	91	50	F92974	F92974	HSPD19063.H
c	28	0.1	47	68	HSM004898	HSM004898	AA104042.Homo.sapi
c	29	0.1	78	50	F29721	F29721	HSPD19771.H
c	30	0.1	83	40	AA988425	AA988425	oc95g12.s
c	31	0.1	90	37	AA708028	AA708028	z65e601.s
c	32	0.1	94	43	AT203131	AT203131	GR34b09.x
c	33	0.1	103	74	AA176421	AA176421	OYO-CT001
c	34	0.1	57	45	AT1434770	AT1434770	t120b06.H
c	35	0.1	69	50	AT687854	AT687854	1P89e03.x
c	36	0.1	83	28	AA117122	AA117122	nm05e09.r
c	37	0.1	89	48	AT559865	AT559865	tc32e12.X
c	38	0.1	91	46	AT453508	AT453508	tj23c08.X
c	39	0.1	105	30	AA216932	AA216932	mv75e12.r
c	40	0.1	54	36	AA603114	AA603114	nc05306.s
c	41	0.1	55	45	AA053911	AA053911	AA053911
c	42	0.1	58	42	AT102195	AT102195	q69e604.x
c	43	0.1	72	62	AT1914923	AT1914923	tt27e08.X
c	44	0.1	73	70	AA155336	AA155336	xx944f09.x
c	45	0.1	74	28	AA078709	AA078709	7T0E11.C

ALIGNMENTS

RESULT 1
T77382/c
LOCUS
DEFINITION
T77382 103 bp mRNA
ydr2h12.c.1 Soares fetal liver spleen INF15 Homo sapiens CDNA clone
IMAGE:118831 5' similar to contains Alu repetitive element// mRNA
sequence.
T77382
T77382.1 GI:694585
EST.
human.
SOURCE
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE
1 (bases 1 to 103)
Hillier,L., Clark,N., Dubuque,T., Elliston,R., Hawkins,M.,
Hollman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
Trevaaskis,E., Waterson,R., Williamson,A., Wohldmann,P. and
Wilson,R.
TITLE
The Washu-Merck EST Project
JOURNAL
Unpublished (1995)
COMMENT
Other_ESTs: ydr2h12.s1
Contact: Wilson RK
Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel.: 314 286 1800
Fax: 314 286 1810
Email: estewatson.wustl.edu
Insert Size: 943
Source: IMAGE Consortium, LBNL This clone is available royalty-free
through LBNL: contact the IMAGE Consortium (info@image.lbnl.gov)
for further information. Putative full length read
Insert Length: 943 Std Error: 0.00
Seq Primer: M13R1
High quality sequence stop: 109.
Location/Qualifiers
1..103

FEATURES

```

/organism="Homo sapiens"
/db_xref="GDB:469448"
/db_xref="taxon:9606"
/clone_image="IMAGE:113831"
/clone_lib="Soares fetal liver spleen mRNAs"
/sex="male"
/dev_stage="20 week post conception fetus"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: Liver and Spleen; Vector: pRTT3D (Pharmacia)
with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
1st strand cDNA was primed with a Pac I - oligo(dT) primer
5' AACGTGAAGATTAATTAAAGATCTTTTTTTTTTTTTC 3';
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Pac I and cloned into the Pac I
and Eco RI sites of the modified pRTT3 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M.Fatima Bonaldo."

```

BASE COUNT
ORIGIN

Query Match	0.3%;	Score 80;	DB 21;	Length 103;
Best Local Similarity	100.0%;	Prod. No. 6.7e-16;		
Matches	80;	Conservative	0;	Mismatches 0;
			Indels	0;
			Gaps	0;
QY	6077	TAGAGCTCTACCTCTACGCAACCTCCCTCCTATATTAAGATGATTCCTGCTCA	6136	
Db	103	TAGAGTCTACCTCTACCTGCAACCTCCCTCCTATATTAAGATGATTCCTGCTCA	44	
QY	6137	GCCTCCGAGTAGCTGGGAC	6156	
Db	43	GCCTCCGAGTAGCTGGGAC	24	

RESULT 2

LOCUS R08388

DEFINITION yf19d03.r1 Soares fetal liver spleen INF18s Homo sapiens cDNA clone IMAGE:127301 5' similar to gp:MK0750_cds1 HISTONE H2B (HUMAN);,

ACCESSION R08388

VERSION R08388

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 95)

AUTHORS Hillier,L., Clark,N., Dubuque,T., Eilston,K., Hawkins,M., Parsons,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Hootson,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevasaki,E., Waterston,R., Williamson,A., Woldmann,P. and Wilson,R.

TITLE The Washb-Merck EST Project

JOURNAL Unpublished (1995)

COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810

Email: est@watson.wustl.edu

Insert Size: 727

High quality sequence stops: 68 Source: IMAGE Consortium, LNL This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.lnl.gov) for further information.

Insert Length: 727 Std Error: 0.00

Seq primer: M13Rpi

High quality sequence stop: 68.

Location/Qualifiers

1. 95

/organism="Homo sapiens"

/db_xref="GDB:479462"

/db_xref="taxon:9606"

/clone="IMAGE:127301"

/clone_lib="Soares fetal liver spleen INTLS"

/sex="male"

/dev_stage="20 week-post conception fetus"

/lab_host="DH10B (ampicillin resistant)"

/note="Organ: liver and spleen; Vector: pT73D (Pharmacia) with a modified polylinker; Site_1: Pac I; Site_2: Eco RI; 1st strand cDNA was primed with a Pac I - oligo (dT) primer (5' AACTGAGAGATTAATTAAGATCTTTTATTTTATTTT 3'), double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Pac I and cloned into the Pac I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization. Library constructed by Benito Soares and M.Fatima Bonaldo."

BASE COUNT 31 a 23 c 26 g 12 t 3 others

ORIGIN

Query Match

Best Local Similarity 100.0%; Pred. No. 1.4e-13;

Matches 73; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10316 TGTTTAAGATGCTGAGCAGCAAGTCTGCTCCGCCGAGAGAGGCTCCAGAG 10375

Db 1 TGTTTAAGATGCTGAGCAGCAAGTCTGCTCCGCCGAGAGAGGCTCCAGAG 60

OY 10376 GCAGTGACCAAG 10388

Db 61 GCAGTGACCAAG 73

RESULT 3

AA500598/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

The WASHU-HMT Mouse EST Project
Unpublished (1996)
On Sep 12, 1996 this sequence version replaced g1:1393909.
Contact: Maria M/Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LNL; contact the

IMAGE Consortium (info@image.lnl.gov) for further information.

MGI:531682

Seq primer: -28m3 rev1 ET from Amersham

High quality sequence stop: 79.

Location/Qualifiers

1. 89

/organism="Mus musculus"

/strain="NIH/Swiss"

/db_xref="taxon:10090"

/clone="IMAGE:919466"

/clone_lib="Stratagene mouse heart (#937316)"

/sex="pooled"

/tissue_type="heart"

/dev_stage="13 day embryos"

/lab_host="SOLR (Xanadyn resistant)"

/note="Organ: heart; Vector: pBluescript SK-; Site_1: EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer: oligo dt. 93 pooled NIH/Swiss 13 day embryo hearts. Average insert size: 1.0 kb. Uni-ZAP XR Vector: -5' adaptor sequence: 5' GAATTCGACGACG 3' -3' adaptor sequence: 5' CTCGAGTTTATTTTATTTT 3'"

BASE COUNT 6 a 8 c 37 g 38 t

ORIGIN

Query Match

Best Local Similarity 100.0%; Pred. No. 0.018;

Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CACACACACACACACACACACACACACACACACAAAT 39

Db 62 CACACACACACACACACACACACACACACACAAAT 24

RESULT 4

T52775

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Generation and analysis of 280,000 human expressed sequence tags
Genome Res. 6 (9), 807-828 (1996)
97044478
Other ESTs: ya79408.s1
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Insert Size: 1013
High quality sequence stops: 82 Source: IMAGE Consortium, LNL This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.lnl.gov) for further information.
Insert Length: 1013 Std Error: 0.00
Seq primer: M13Rpi
High quality sequence stop: 82.
Location/Qualifiers

source
1. 95
/organism="Homo sapiens"
/db_xref="GDB:504144"
/db_xref="taxon:9606"
/clone="IMAGE:67887"
/clone_1b="Stratagene ovary (#937217)"
/sex="female"
/dev_stage="49 year old"
/lab_host="SOLR cells (kanamycin resistant)"
/note="Organ: ovary; Vector: Bluescript SK; Site_1: EcoRI;
Site_2: XhoI; Cloned unidirectionally. Primer: Oligo dt.
Total ovary tissue, normal, caucasian. Average insert
size: 0.8 kb; Uni-ZAP XR Vector; -5' adaptor sequence: 5'
GAATTCGCGACGAG 3' -3' adaptor sequence: 5'
CTCGAGTTTCTTTTCTTTT 3' "

BASE COUNT
19 a 32 c 20 g 24 t

ORIGIN

Query Match 0.1%; Score 38; DB 20; Length 95;
Best Local Similarity 100.0%; Pred. No. 0.037;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13817 CAGTGGCAGCATCTGGCTCAGTCAACCTCTGCTCC 13854
|||||
Db 27 CAGTGGCAGCATCTGGCTCAGTCAACCTCTGCTCC 64

RESULT 5
AQ006534/c 101 bp DNA GSS 26-JUN-1998
LOCUS AQ006534
DEFINITION genomic survey sequence.
ACCESSION AQ006534
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 101)
Adams, M.D., Rounsley, S.D., Field, C.E., Bass, S., Linher, K.,
Golden, K., Berry, K., Granger, D., Suh, E., Wible, C., Kim, U.-J.,
Shizuya, H., Simon, M., and Venter, J.C.
Use of a human BAC End Sequence Database for Sequence-Ready Map
Building
Unpublished (1997)
Other_GSS: CIT-HSP-2054B16.TR.1 CIT-HSP-2054B16.TR
Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tdb/humgen/Bac_end_search/Bac_end_search.html
Seq primer: M13-21;
Class: BAC ends.

JOURNAL
TITLE
AUTHORS
REFERENCE
COMMENT

FEATURES
source
1. 101
/organism="Homo sapiens"
/db_xref="GDB:705763"
/db_xref="taxon:9606"
/clone="2054B16"
/clone_1b="CIT-HSP"
/sex="Male"
/cell_type="Sperm"
/note="Vector: pBelBAC11; Site_1: HindIII; Site_2:
HindIII"
BASE COUNT
25 a 12 c 24 g 40 t

ORIGIN

Query Match 0.1%; Score 38; DB 90; Length 101;
Best Local Similarity 100.0%; Pred. No. 0.036;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CACACACACACACACACACACACACACACACACAA 38
|||||
Db 98 CACACACACACACACACACACACACACACACAAA 61

RESULT 6
AA935225/c 54 bp mRNA EST 07-JUL-1998
LOCUS AA935225
DEFINITION o06b11.s1 NCI-CGAP_G4 Homo sapiens cDNA clone IMAGE:1571133 3'
Similar to gb:144812 RETINOBLASTOMA-ASSOCIATED PROTEIN-LIKE 107 KD
HOMOLOG (HUMAN); contains LTR2.t2 LTR2 repetitive element ;, mRNA
sequence.
ACCESSION AA935225.1 GI:3092382
VERSION AA935225
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 54)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
On Jan 14, 1998 this sequence version replaced gi:1797423.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium LTRU at:
www.bio.lnl.gov/bbrp/image/image.html

JOURNAL
TITLE
AUTHORS
REFERENCE
COMMENT

FEATURES
source
1. 54
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1571133"
/clone_1b="NCI-CGAP_G4"
/tissue="pooled germ cell tumors"
/lab_host="DH10B"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; 1st strand cDNA was prepared from 3 pooled
germ cell tumors, and was primed with a Not I -
oligo(dT) primer. Double-stranded cDNA was ligated to Eco
RI adaptors (Pharmacia), digested with Not I and cloned
into the Not I and Eco RI sites of the modified pT73
vector. Library is normalized. Library was constructed by
Bento Soares and M. Fatima Bonaldo."

BASE COUNT
8 a 12 c 17 g 17 t

ORIGIN

Query Match 0.1%; Score 37; DB 40; Length 54;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24632 CCATGCACTCGAGCTGGGACACAGCAAACTC 24658
|||||

Db 42 CCATTGCATCTCAGCTGGCGACACAGACAAAATC 6

RESULT 7
HSM010004/c
ID HSM010004 standard; RNA; EST; 58 BP.
XX
AC AL045154;
XX
SV AL045154.1
XX
DT 12-MAR-1999 (Rel. 59, Created)
DT 12-MAR-1999 (Rel. 59, last updated, Version 1)
XX
DE Homo sapiens mRNA; EST DKFZp434P234_r1 (from clone
DE DKFZp434P234)
XX
KM EST; expressed sequence tag.
XX
OS Homo sapiens (human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria;
OC Primates; Catarrhini; Hominoidea; Homo.
XX
RN [1]
RP 1-58
RA Wambutt R., Heubner D., Mewes W., Gassenhuber J., Wiemann S.;
RT Submitted (12-MAR-1999) to the EMBL/GenBank/DBJ databases.
RL MIPs, Am Klopferspitz 18a D-82152 Martinsried, GERMANY
XX
CC Clone from S. Wiemann, sequenced by AGOWA within the cDNA
CC sequencing consortium of the German Genome Project
CC s1 sequence also available
CC This clone is available at the RZPD in Berlin
CC Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
CC Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de
XX
FH Key Location/Qualifiers
FH 1. 58
FH source
FT /db_xref="taxon:9606"
FT /organism="Homo sapiens"
FT /clone="DKFZp434P234"
FT /clone_lib="434 (synonym: htes3). Vector pSPORT1; host
FT DH10B; sites NotI, SalI
FT /dev_stage="adult"
FT /tissue_type="testis"
XX
SQ Sequence 58 BP; 13 A; 2 C; 21 G; 22 T; 0 other;

Query Match 0.1%; Score 37; DB 68; Length 58;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 ACACACACACACACACACACACACACACACACAAA 38
Db 58 ACACACACACACACACACACACACACACACACAAA 22

RESULT 8
LOCUS F24490 80 bp mRNA EST 13-MAY-1999
DEFINITION HSPD10834 HM3 Homo sapiens cDNA clone s400013A06, mRNA sequence.
ACCESSION F24490
VERSION F24490.1 GI:4810116
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE
1 (bases 1 to 80)
Lanfranchi, G., Muraro, T., Caldara, F., Pacchioni, B., Pallavicini, A.,

TITLE
JOURNAL
MEDLINE
COMMENT
FEATURES
source
1. 80
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="s400013A06"
/clone_lib="HM3"
/sex="female"
/tissue_type="pectoral muscle (after mastectomy)"
/note="Vector: pCDNAII (Invitrogen); Site_1: BstXI;
Site_2: NotI. The library was constructed by G.
Lanfranchi. This library is not subtracted nor normalized.
The first strand cDNA was primed with a biotinylated
oligo-dT-NotI primer
(5'-biotin-AACCGGCTCGACGCGCGCTTTTCTTTTCTTTT-3'). The
35 cDNA was sonicated and size-selected in the range
350-550 bp. The 3' specific fragments were selected by
streptavidin coated magnetic beads, ligated to
non-palindromic BstXI adapters, NotI digested and
directionally cloned into BstXI-NotI cut pCDNAII vector."

BASE COUNT
ORIGIN
18 a 21 c 19 g 22 t

Query Match 0.1%; Score 37; DB 50; Length 80;
Best Local Similarity 100.0%; Pred. No. 0.087;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 23005 GGGCGAGTGGCTCATGCTGTAAATCCACACTTTG 23041
Db 37 GGGCGAGTGGCTCATGCTGTAAATCCACACTTTG 73

RESULT 9
LOCUS AA598741/c
DEFINITION
ACCESSION AA598741
VERSION AA598741.1 GI:2432413
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE
1 (bases 1 to 89)
Hillier, L., Allen, M., Bowles, L., Dubuque, T., Giesel, G., Jost, S.,
Kizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M.,
Martin, J., Moore, B., Schellenberg, R., Steptoe, M., Tan, F.,
Thelings, B., White, Y., Wyllie, T., Waterston, R. and Wilson, R.
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LNL; contact the

TITLE
JOURNAL
MEDLINE
COMMENT
FEATURES
source
1. 80
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="s400013A06"
/clone_lib="HM3"
/sex="female"
/tissue_type="pectoral muscle (after mastectomy)"
/note="Vector: pCDNAII (Invitrogen); Site_1: BstXI;
Site_2: NotI. The library was constructed by G.
Lanfranchi. This library is not subtracted nor normalized.
The first strand cDNA was primed with a biotinylated
oligo-dT-NotI primer
(5'-biotin-AACCGGCTCGACGCGCGCTTTTCTTTTCTTTT-3'). The
35 cDNA was sonicated and size-selected in the range
350-550 bp. The 3' specific fragments were selected by
streptavidin coated magnetic beads, ligated to
non-palindromic BstXI adapters, NotI digested and
directionally cloned into BstXI-NotI cut pCDNAII vector."

VERSION AI262095.1 GI:3870298
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 80)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT On Jan 14, 1998 this sequence version replaced gi:1798653.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Christopher Moskalko, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA sequencing by: Washington University Genome Sequencing Center
clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/BLNI at:
www.bio.linn.gov/dbtrp/image/image.html

FEATURES
source
Seq primer: -40RP from GIBCO.
Location/Qualifiers
1..80
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="2028200"
/clone_lib="NCI-CGAP_K1d11"
/lab_host="DH10B"
/note="Organ: kidney; Vector: pT7AD-Pac (Pharmacia) with
a modified polylinker; Site_1: Not I; Site_2: Eco RI;
Plasmid DNA from the normalized library NCI-CGAP_K1d3 was
prepared, and ss circles were made in vitro. Following HAP
purification, this DNA was used as tracer in a subtractive
hybridization reaction. The driver was PCR-amplified cDNAs
from a pool of 5,000 clones made from the same library
(clones 132376-132391, 145607-145675, and
150052-150285). Subtraction by Bento Soares and M.
Fatima Bonaldo."

BASE COUNT 13 a 25 c 24 g 18 t
ORIGIN

Query Match 0.1%; Score 36; DB 44; Length 80;
Best Local Similarity 100.0%; Pred. No. 0.19;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 6248 CCCAGCTGAGTGCAGTGCACGATCTGGCTCAC 6283
|||||
Db 24 CCCAGCTGAGTGCAGTGCACGATCTGGCTCAC 59

RESULT 13
F34634 84 bp mRNA EST 13-MAY-1999
LOCUS HSPD29754 HM3 Homo sapiens cDNA clone SH3-000002-1/F06, mRNA
DEFINITION sequence.
ACCESSION F34634
VERSION F34634.1 GI:4820260
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 84)
AUTHORS Lanfanchi, G., Muraro, T., Caldara, F., Paschioni, B., Pallavicini, A.,
Pandofo, D., Toppo, S., Trevisan, S., Scarso, S. and Valle, G.
TITLE Identification of 4370 expressed sequence tags from a
3'-end-specific cDNA library of human skeletal muscle by DNA
sequencing and filter hybridization

JOURNAL Genome Res. 6 (1), 35-42 (1996)
MEDLINE 96276048
COMMENT On Jun 5, 1998 this sequence version replaced gi:3188981.
Contact: Valle G.
CIRI Biotechnology Centre
University of Padua
Via Trieste 75, 35121 Padua, Italy
ABI Chromatograms and other information are available on WWW at
<http://grup.bio.unipd.it>.

FEATURES
source
Location/Qualifiers
1..84
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="SH3-000002-1/F06"
/clone_lib="HM3"
/sex="female"
/tissue_type="pectoral muscle (after mastectomy)"
/note="Vector: pCDNAII (Invitrogen); Site_1: BstXI;
Site_2: NotI; The library was constructed by G.
Lanfanchi. This library is not subtracted nor normalized.
The first strand cDNA was primed with a biotinylated
oligo-dT-NotI primer
(5'-biotin-AACCGCGCTCGAGCGCCGCTTTTCTTTTCTTTT-3'). The
ds cDNA was sonicated and size-selected in the range
350-550 bp. The 3' specific fragments were selected by
streptavidin coated magnetic beads, ligated to
non-palindromic BstXI adapters, NotI digested and
directionally cloned into BstXI-NotI cut pCDNAII vector."

BASE COUNT 18 a 17 c 23 g 25 t
ORIGIN

Query Match 0.1%; Score 36; DB 50; Length 84;
Best Local Similarity 100.0%; Pred. No. 0.18;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 23006 GGGCGAGTGGCTCATGCTGATCCAGCACTTGG 23041
|||||
Db 42 GGGCGAGTGGCTCATGCTGATCCAGCACTTGG 77

RESULT 14
AA078242/c 89 bp mRNA EST 24-SEP-1999
LOCUS 7H16G05 Chromosome 7 Hela cDNA library Homo sapiens cDNA clone
DEFINITION 7H16G05, mRNA sequence.
ACCESSION AA078242
VERSION AA078242.1 GI:1837716
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 89)
AUTHORS Touchman, J.W., Bouffard, G.G., Weintraub, L.A., Idol, J.R., Wang, L.,
Robbins, C.M., Nussbaum, J.C., Lovett, M. and Green, E.D.
TITLE 2006 expressed-sequence tags derived from human chromosome
7-enriched cDNA libraries
JOURNAL Genome Res. 7 (3), 281-292 (1997)
MEDLINE 97228905
COMMENT On Sep 12, 1996 this sequence version replaced gi:1292359.
Contact: Eric D. Green
Genome Technology Branch
National Human Genome Research Institute/NIH
49 Convent Dr., MSC4431, Building 49, Room 2A08, Bethesda, MD 20892
Tel: 3014020201
Fax: 3014024735
Email: egreen@hgrl.nih.gov
Plate: 16 row: G column: 05
Seq primer: -21M13 (ABI).
Location/Qualifiers
1..89
/organism="Homo sapiens"

```

BASE COUNT
ORIGIN
0 a
      0 c
        45 g
          44 t
            /db_xref="taxon:5606"
              /clone="7H16G05"
                /clone_lib="Chromosome 7 HeLa cDNA Library"
                  /sex="female"
                    /cell_line="HeLa cell line: ATCC"
                      /lab_host="E. coli strain DH5 alpha"
                        /note="Vector: pAMP10: cDNA was generated from cytoplasmic
                          RNA using a mixture of random DNA hexamers and oligo(dT)
                          from this pool of cDNA. human chromosome 7-enriched cDNA
                          was isolated by direct cDNA selection using chromosome 7
                          genomic DNA (cosmids). The resulting direct-selected cDNA
                          was cloned into a plasmid vector using a non-directional
                          uracil DNA glycosylase (UDG)-mediated cloning strategy."

```

	Query Match	Similarity	0.1%;	Score	36;	DB	28;	Length	89;	
	Best Local	Similarity	100.0%;	Pred.	No.	0.17;				
	Matches	36;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
Oy	1	CACACACACACACACACACACACACACA	36							
Db	89	CACACACACACACACACACACACACACA	54							

RESULT	15	
AA598741		
LOCUS		
DEFINITION	AA598741 89 bp mRNA	19-DEC-1997
	ae44912.s1 StrataGene lung carcinoma 937218 Homo sapiens CDNA clone	saplen CDNA clone
	IMAGE:950207.3 similar to contains Alu repetitive element; mRNA	
	sequence.	
ACCESSION	AA598741	
VERSION	AA598741.1	GI:2432413
KEYWORDS	EST.	
SOURCE	human.	

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Dukerjy, T., Metacoa: Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Carnivori; Homnidae; Homo.
1 (bases 1 to 89)
Hillier, L., Allen, M., Bowles, L., Dubouque, T., Geisel, G., Jost, S., Krisman, D., Kuchaba, T., Lacy, M., Le, N., Lemco, G., Marra, M., Martin, J., Moore, B., Schellenberg, K., Stepto, M., Tan, F., Theising, B., White, Y., Wylie, T., Waterston, R. and Wilson, R.
Wasnu-NCI human EST Project
Unpublished (1997)
On May 18, 1995 this sequence version replaced gi:811447.

Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: estewatson.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further
information.
Seq primer: -40ml3 fwd. ET from Amersham.

```

FEATURES
source
Location/Qualifiers
1. 89
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image:950278-
/clone_lib="Stratagene lung Carcinoma 937218"
/tissue_type="lung carcinoma"
/cell_line="NCI-H69"
/dev_stage="cell line NCI-H69"
/lab_host="SOLR (kanamycin resistant)"
/note="Organ: lung; Vector: plasmidirect SK-; Site:1: EcoRI; Site:2: XhoI; Cloned unidirectionally. Primer: Oligo dT. Small cell carcinoma cell line NCI-H69. Average insert size: 1.0 kb. Inl-ZAP XR Vector. 5' adaptor sequence: 5' GAATTCGCGACGAG 3' -3' adaptor sequence: 5' CTCGAGTTTCTTTTCTTTT 3'"
13 a 22 c 25 g 29 t
BASE COUNT

```

```

ORIGIN
Query Match                                0.1%; Score 36; DB 36; Length 89;
Best local Similarity 100.0%; Pred.No. 0.17;
Matches   36; Conservative    0; Mismatches 0; Indels      0; Gaps       0;
QY  12199 CCCAGGCTGGAGTGCATGTGCGCGCATCTTGACTCAC 12234
      |||.....|||.....|
Db   28    CCCAGGCTGGAGTGCATGTGCGCGCATCTTGACTCAC 63

```

Search completed: March 29, 2000, 06:25:57
Job time: 11411 sec

Mon Apr 3 08:24:18 2000

us-08-852-495c-1_copy_1_30000.rst

Page 9



•

•

•

05 Bos taurus.
PN MO9213102-A.
PD 06-AUG-1992.
PF 15-JAN-1992; U00340.
PR 15-JAN-1991; US-642342.
PA (GENM.) GENMARK.
PI Georges M. Massey JM;
DR WPI: 92-284684/34.
PT polymorphic bovine DNA markers - used in genetic identification,
PR gene mapping, and selective breeding
PS Table 7; page 269; 517pp; English.
CC The sequence is that of a bovine microsatellite sequence obt. by
CC screening a library of bovine Mhoi DNA fragments of between
CC 250 and 500 bp with an (AC)₁₅ and a (TC)₁₅ oligonucleotide probe.
CC One out of 50 clones cross-hybridised. Assuming independent
CC distribution of microsatellites and Mhoi sites, the frequency of
CC (TC)₁₅ >9 microsatellites in the bovine genome is estimated at >100,
CC 000. The sequence information for ca. 230 such bovine microsatellites
CC is summarised in the specification and indexed herein (see below).
CC The sequences upstream and downstream of the microsatellite sequence
CC were used to generate the required PCR primers for in vitro
CC amplification of the corresp. microsatellites (using the program
CC OPLPRIM). The microsatellites may be used to identify individuals,
CC for parentage testing, and in the genetic mapping of economic trait
CC loci, or genes involved in the determinism of economically important
CC traits esp. in cattle, to allow selective breeding.
CC See also Q33501-34437.
EO Sequence 70 BP; 1 A; 1 C; 33 G; 35 T;

```

Query Match      0.2% Score 41; DB 1; Length 70;
Best Local Similarity 100.0%; Pred. No. 0.00018;
Matches    41; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```

RESULT 3
ID V19044/c
AC V19044:
DT 28-JUL-1998 (first entry)
DE Alu PCR primer 1.
KW PCR; primer; amplification; Alu repeat sequence; vector;
KW circular yeast artificial chromosome; YAC; ss.
OS synthetic.
OS Saccharomyces sp.
PN MO9801573-A1.
PD 15-JAN-1998.
PF 09-JUL-1996; U11478.
PR 09-JUL-1996; MO-U11478.
PA (USSR) US DEPT HEALTH & HUMAN SERVICES.
PI Koupina NY, Iarionov VL, Perkins EL, Resnick MA;
DR WFL; 98-110234/10.
PT Preparation of yeast artificial chromosomes - by *in vivo*
PT recombination using vector comprising yeast centromere, marker,
PT yeast telomere and nucleic acid for recombination
PS Example 1; Page 45; 117pp; English.
CC This is the nucleotide sequence for the PCR primer used in the
CC amplification of the Alu repeat sequence, which is used to
CC demonstrate the processes described in the invention. It involves
CC the creation and use of circular yeast artificial chromosome (YAC)
CC to selectively clone specific nucleic acids from a background of
CC mixed nucleic acids by introducing the vector(s) into *E. coli* cells.
CC They can be used to rapidly isolate human DNA where only a part of the
CC sequence of DNA is known. Using the methods large fragments of DNA can
CC be easily cloned and analysed.
SO Sequence 40 Bp. 7 A; 12 C; 13 G; 8 T;

Query Match 0.28; Score 40; DB 1; Length 40;

```

Best Local Similarity 100.0%; Pred. No. 0.00042; Indels 0; Gaps 0;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 8640 TCGGCTCCCAAGTGTGGATTACAGCGGTGAGCCACC 8679
      |||||||
Db 40 TCGGCTCCCAAGTGTGGATTACAGCGGTGAGCCACC 1

RESULT 4
T65714/C
ID AC T65714 standard; DNA; 39 BP.
DT 17-JUN-1997 (first entry)
DE Repeat sequence from polymorphic marker Med12.
KW Polymorphism; repeat sequence; genetic marker; primer; amplification;
  PCR; polymerase chain reaction; paternity; human; pedigree;
  linkage analysis; genetic disease; animal; plant; breeding; locus;
  hybridisation; chromosome; ds.
KW Homo sapiens.
PN US5582979-A.
PD 10-DEC-1996.
PR 21-APR-1989; 341562.
PR 21-APR-1989; US-341562.
PR 05-SEP-1991; US-754351.
PR 04-APR-1994; US-222177.
PA (MARS-) MARSHFIELD CLINIC.
PI Weber JL;
DR 97-042239/04.
PT Detection of polymorphic genetic markers of the form
  (dC-dA)n - using novel nucleic acid mole. as primers
  (dC-dA)n(dG-dT)n - using novel nucleic acid mole. as primers
  Disclosure; Column 9-10; 186pp; English.
PS The invention relates to the isolation of polymorphic repeat sequences
  CC having the sequence (dC-dA)n.(dG-dT)n which can be used as genetic
  CC markers. Primers based on these sequences can be used to detect these
  CC repeats, especially for use in e.g. paternity or maternity testing,
  CC human genetic analysis such as linkage analysis of genetic disease,
  CC commercial animal or plant breeding or pedigree analysis. Clones
  CC containing the repeat sequences were isolated by hybridisation of
  CC chromosome-specific phage libraries with a synthetic poly(dC-dA).(dG-dT)
  CC probe. Over 100 repeat blocks were isolated. The inserts from the clones
  CC were amplified by primers T65798-T66047. Those clones where the repeat
  CC sequence has been determined are shown in T65704-797. This repeat
  CC sequence is from the marker clone K611 which contains the repeat
  CC sequence having the formula (AC)11AT(AC)8A.
  CC Sequence 39 BP; 20 A; 18 C; 0 G; 1 T;

Query Match 0.2%; Score 39; DB 1; Length 39;
Best Local Similarity 100.0%; Pred. No. 0.00088;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 7522 TGTGTGTGTGTGTATGTGTGTGTGTGTGTGTGTGT 7560
      |||||||
Db 39 TGTGTGTGTGTGTGTATGTGTGTGTGTGTGTGTGTGT 1

RESULT 5
O34003
ID AC O34003 standard; DNA; 50 BP.
DT 02-FEB-1993 (first entry)
DE Microsatellite sequence from clone TG1A400.
KW PCR; selection; primers; OPIPPIW; breeding; cattle; parentage;
  KW genetic mapping; traits; amplification; ss.
KW Bos taurus.
PN WO9213102-A.
PD 06-AUG-1993.
PR 15-JUN-1991; U00340.
PR 15-JUN-1991; US-642342.
PA (GENM-) GENMARK.
PI George M, Massey JM;
DR 92-28664/34.
PT Polymorphic bovine DNA markers - used in genetic identification,

```

PT Polymorphic bovine DNA markers - used in genetic identification,

DT 14-SEP-1998 (first entry)
 DE Oligonucleotide sequence of the specification.
 KW Genotype: phenotype; molecular evolutionary engineering;
 KW functional biopolymer; virus; ss.
 OS Synthetic.
 PN MO9816636-A1.
 PD 23-APR-1998.
 PR 17-OCT-1998; J03766.
 PA (MITU) MITSUBISHI CHEM. CORP.
 PI Fusimi Y, Miyamoto E, Nemoto N, Yanagawa H;
 DR WPI; 98-261039/23.
 PT Virus containing nucleic acid and protein sections - for use in
 PT modification and creation of functional bio:polymers such as
 PT enzymes, antibodies and ribozyme(s)
 PS Example 1; Page 44; 68pp; Japanese.
 CC The present sequence is used in the course of the invention. The
 CC specification describes a molecule for bringing together genotype with
 CC phenotype (in vitro virus). The molecule contains a nucleic acid
 CC fragment having a base sequence corresponding to a genotype, covalently
 CC bonded to a protein fragment containing a protein participating in
 CC phenotype expression, the 3'-end of the nucleic acid part being bonded
 CC to the C-terminus of the protein part via a puromycin moiety. The nucleic
 CC acid fragment preferably consists of RNA corresponding to the gene (free
 CC from a termination codon), a spacer (such as two-stranded DNA), a peptide
 CC adapter (such as a DNA-tRNA hybrid containing an anticodon corresponding
 CC to the terminator codon of the gene) and a puromycin-containing cap
 CC capable of binding to an amino acid residue. Translation of the virus is
 CC performed in a non-cellular (preferably ribosomal) system (e.g. E. coli
 CC ribosome), the protein synthesised by the translation attaching to the
 CC puromycin cap, resulting in the complete in vitro virus structure. The
 CC method is used in molecular evolutionary engineering to optimise function
 CC of a functional biopolymer such as an enzyme, antibody or ribozyme, or
 CC to generate new functionality. The virus can be used for the functional
 CC optimisation of nucleic acid or protein sequences.
 SQ Sequence 80 BP; 7 A; 10 C; 1 G; 62 T;

Query Match 0.2%; Score 33; DB 1; Length 80;
 Best Local Similarity 100.0%; Pred. No. 0.064;
 Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 4426 AATAATTTTTTTTTTTTTTTTTTTTTTTT 4458
 DB 6 AATAATTTTTTTTTTTTTTTTTTTTTTTT 38

Search completed: March 29, 2000, 17:55:02
 Job time: 30700 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 29, 2000, 12:57:56 ; Search time 229.02 Seconds
(without alignments)
10455.780 Million cell updates/sec

Title: US-08-852-495c-1_COPY_25000_45000

Perfect score: 20001
Sequence: 1 AGTCAGTAGTGCAGACCTC.....CAAGTTAGACATTTT 20001

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 214294 seqs, 59861574 residues

Word size: 0

Total number of hits satisfying chosen parameters: 349160

Minimum DB seq length: 8

Maximum DB seq length: 105

Post-processing: Listing first 45 summaries

Database:

Issued_Patents_NA: *
1: /cgn2_6/ptodata/1/lna/3A.COMB.seq:*
2: /cgn2_6/ptodata/1/lna/3B.COMB.seq:*
3: /cgn2_6/ptodata/1/lna/3C.COMB.seq:*
4: /cgn2_6/ptodata/1/lna/3D.COMB.seq:*
5: /cgn2_6/ptodata/1/lna/6.COMB.seq:*
6: /cgn2_6/ptodata/1/lna/PTU59.COMB.seq:*
7: /cgn2_6/ptodata/1/lna/backfilssl.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	48	0.2	56	1	US-08-222-177A-149 Sequence 149, App
2	43	0.2	54	2	US-08-469-802B-27 Sequence 27, App
3	43	0.2	54	3	US-08-267-803B-45 Sequence 45, App
4	39	0.2	39	1	US-08-222-177A-86 Sequence 86, App
5	39	0.2	39	2	US-08-632-673B-1 Sequence 1, App
6	38	0.2	60	3	US-08-454-557C-57 Sequence 57, App
7	38	0.2	60	4	US-08-340-426D-57 Sequence 57, App
8	38	0.2	60	4	US-08-450-673C-57 Sequence 57, App
9	38	0.2	60	6	PCT-US95-17111A-57 Sequence 57, App
10	38	0.2	66	2	US-08-469-802B-29 Sequence 29, App
11	38	0.2	66	3	US-08-267-803B-47 Sequence 47, App
12	38	0.2	76	4	US-08-454-557C-69 Sequence 69, App
13	38	0.2	76	4	US-08-340-426D-69 Sequence 69, App
14	38	0.2	76	4	US-08-450-673C-69 Sequence 69, App
15	38	0.2	76	6	PCT-US95-17111A-69 Sequence 69, App
16	36	0.2	40	1	US-08-222-177A-152 Sequence 152, App
17	35	0.2	45	1	US-08-222-177A-157 Sequence 157, App
18	35	0.2	85	3	US-08-454-557C-92 Sequence 92, App
19	35	0.2	85	4	US-08-340-426D-92 Sequence 92, App
20	35	0.2	85	4	US-08-450-673C-92 Sequence 92, App
21	35	0.2	85	6	PCT-US95-17111A-92 Sequence 92, App
22	31	0.2	80	2	US-07-920-281C-25 Sequence 25, App
23	31	0.2	90	2	US-08-677-944-1 Sequence 1, App
24	30	0.1	37	1	US-08-113-646A-44 Sequence 44, App
25	30	0.1	43	1	US-08-222-177A-370 Sequence 370, App
26	30	0.1	56	4	US-08-776-944-9 Sequence 9, App

27	30	0.1	70	4	US-08-776-944-12 Sequence 12, App
28	30	0.1	75	4	US-08-776-944-13 Sequence 13, App
29	30	0.1	90	2	US-08-677-944-2 Sequence 2, App
30	29	0.1	30	2	US-08-433-505-9 Sequence 9, App
31	29	0.1	40	1	US-08-222-177A-175 Sequence 175, App
32	29	0.1	40	4	US-08-771-624B-1 Sequence 1, App
33	29	0.1	40	4	US-08-771-624B-10 Sequence 10, App
34	29	0.1	40	4	US-08-440-209-4 Sequence 4, App
35	29	0.1	42	1	US-07-875-167-2 Sequence 2, App
36	29	0.1	42	1	US-08-222-177A-53 Sequence 53, App
37	29	0.1	42	1	US-08-287-164-2 Sequence 2, App
38	29	0.1	44	1	US-08-113-646A-40 Sequence 40, App
39	29	0.1	44	2	US-08-664-596B-9 Sequence 9, App
40	29	0.1	47	4	US-08-778-494B-114 Sequence 114, App
41	29	0.1	50	1	US-08-233-609-5 Sequence 5, App
42	29	0.1	50	1	US-08-381-572-20 Sequence 20, App
43	29	0.1	50	1	US-08-444-083-5 Sequence 5, App
44	29	0.1	50	1	US-08-286-304-5 Sequence 5, App
45	29	0.1	50	1	US-08-420-443-1 Sequence 1, App

ALIGNMENTS

RESULT 1
US-08-222-177A-149/C
Sequence 149, Application US/08222177A
Patent No. 5582979
GENERAL INFORMATION:
APPLICANT: Weber, James L.
TITLE OF INVENTION: LENGTH POLYMORPHISMS IN
TITLE OF INVENTION: (dc-da)n.(dc-dt)n SEQUENCES AND METHODS OF USING SAME
NUMBER OF SEQUENCES: 460
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dewitt Ross & Stevens, S.C.
STREET: 8000 Excelsior Drive, Suite 401
CITY: Madison
STATE: Wisconsin
COUNTRY: USA
ZIP: 53717-1914
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/222,177A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/341,562
FILING DATE: 21-APR-1989
ATTORNEY/AGENT INFORMATION:
NAME: Sara, Charles S.
REGISTRATION NUMBER: 30,492
REFERENCE/DOCKET NUMBER: 09865.601
TELECOMMUNICATION INFORMATION:
TELEPHONE: (608) 831-2100
TELEFAX: (608) 831-2106
TELEX:
INFORMATION FOR SEQ ID NO: 149:
SEQUENCE CHARACTERISTICS:
LENGTH: 56 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
IMMEDIATE SOURCE:
CLONE: mid33its
US-08-222-177A-149
Query Match 0.2%; Score 48; DB 1; Length 56;


```

? TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection
? NUMBER OF SEQUENCES: 121
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
? STREET: 1100 New York Avenue, Suite 600
? CITY: Washington
? STATE: D.C.
? COUNTRY: U.S.A.
? ZIP: 20005-3934
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patent Release #1.0, Version #1.25
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/340,426D
? FILING DATE: 14-NOV-1994
? CLASSIFICATION: 435
? ATTORNEY/AGENT INFORMATION:
? NAME: Ludwig, Steven R.
? REGISTRATION NUMBER: 36,203
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (202) 371-2500
? TELEFAX: (202) 371-2540
? INFORMATION FOR SEQ ID NO: 57:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 60 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: both
? TOPOLOGY: both
?
US-08-340-426D-57
?
Query Match          0.2%: Score 38; DB 4; Length 60;
Best Local Similarity 100.0%; Pred. No. 9.2e-05;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12897 GGTGAACCCCTGCTCTACTAAATACAAATTAGC 12934
DB 48 GGTGAACCCCTGCTCTACTAAATACAAATTAGC 11

RESULT 8
US-08-450-673C-57/c
? Sequence 57, Application US/08450673C
? Patent No. 5948868
? GENERAL INFORMATION:
? APPLICANT: de la Monte, Suzanne
? ATTORNEY/AGENT INFORMATION:
? NAME: Ludwig, Steven R.
? REGISTRATION NUMBER: 36,203
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (202) 371-2500
? TELEFAX: (202) 371-2540
? INFORMATION FOR SEQ ID NO: 57:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 60 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: both
? TOPOLOGY: both
?
PCT-US95-17111A-57
?
Query Match          0.2%: Score 38; DB 4; Length 60;
Best Local Similarity 100.0%; Pred. No. 9.2e-05;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12897 GGTGAACCCCTGCTCTACTAAATACAAATTAGC 12934
DB 48 GGTGAACCCCTGCTCTACTAAATACAAATTAGC 11

PCT-US95-17111A-57/c
? Sequence 57, Application PC/TUS9517111A
? GENERAL INFORMATION:
? APPLICANT: de la Monte, Suzanne
? ATTORNEY/AGENT INFORMATION:
? NAME: Ludwig, Steven R.
? REGISTRATION NUMBER: 36,203
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (202) 371-2500
? TELEFAX: (202) 371-2540
? INFORMATION FOR SEQ ID NO: 57:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 60 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: both
? TOPOLOGY: both
?
US-08-450-673C-57
?
Query Match          0.2%: Score 38; DB 6; Length 60;
Best Local Similarity 100.0%; Pred. No. 9.2e-05;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12897 GGTGAACCCCTGCTCTACTAAATACAAATTAGC 12934
DB 48 GGTGAACCCCTGCTCTACTAAATACAAATTAGC 11

```

```

? REFERENCE/DOCKET NUMBER: 0609.3840004
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (202) 371-2600
? TELEFAX: (202) 371-2540
? INFORMATION FOR SEQ ID NO: 57:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 60 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: both
? TOPOLOGY: both
?
US-08-450-673C-57
?
Query Match          0.2%: Score 38; DB 4; Length 60;
Best Local Similarity 100.0%; Pred. No. 9.2e-05;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12897 GGTGAACCCCTGCTCTACTAAATACAAATTAGC 12934
DB 48 GGTGAACCCCTGCTCTACTAAATACAAATTAGC 11

PCT-US95-17111A-57/c
? Sequence 57, Application PC/TUS9517111A
? GENERAL INFORMATION:
? APPLICANT: de la Monte, Suzanne
? ATTORNEY/AGENT INFORMATION:
? NAME: Ludwig, Steven R.
? REGISTRATION NUMBER: 36,203
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (202) 371-2500
? TELEFAX: (202) 371-2540
? INFORMATION FOR SEQ ID NO: 57:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 60 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: both
? TOPOLOGY: both
?
PCT-US95-17111A-57
?
Query Match          0.2%: Score 38; DB 6; Length 60;
Best Local Similarity 100.0%; Pred. No. 9.2e-05;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12897 GGTGAACCCCTGCTCTACTAAATACAAATTAGC 12934
DB 48 GGTGAACCCCTGCTCTACTAAATACAAATTAGC 11

```

```

RESULT 10
US-08-469-802B-29
; Sequence 29, Application US/08469802B
Patent No. 5741645
GENERAL INFORMATION:
APPLICANT: Orr, Harry T.
APPLICANT: Raunum, Laura P.W.
APPLICANT: Chung, Ming-yi
APPLICANT: Zoghbi, Huda Y.
TITLE OF INVENTION: Gene Sequence for Spinocerebellar Ataxia
Patent No. 5741645
TITLE OF INVENTION: Type I and Method for Diagnosis
NUMBER OF SEQUENCES: 47
CORRESPONDENCE ADDRESS:
ADDRESSEE: Mueeling, Raasch, Gebhardt & Schwappach, P.A.
STREET: 119 N. 5741645th Fourth Street, Suite 203
CITY: Minneapolis
STATE: MN
COUNTRY: USA
ZIP: 55401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/469, 802B
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Mueeling, Ann M.
REGISTRATION NUMBER: 33,977
REFERENCE/DOCKET NUMBER: 110.00030101
TELECOMMUNICATION INFORMATION:
TELEPHONE: 612-305-1217
TELEFAX: 612-305-1225
INFORMATION FOR SEQ ID NO: 29:
SEQUENCE CHARACTERISTICS:
LENGTH: 66 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-469-802B-29

Query Match          0.2%; Score 38; DB 2; Length 66;
Best Local Similarity 100.0%; Pred. No. 9.1e-05;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 7513 GTGTGTTGTGTGTGTGTGTGTGTATGTGTGTGT 7550
      |||||||
Db 29 GTGTGTGTGTGTGTGTGTGTGTGTATGTGTGTGTGT 66

RESULT 11
US-08-267-803B-47
; Sequence 47, Application US/08267803B
Patent No. 5834183
GENERAL INFORMATION:
APPLICANT: Orr, Harry T.
APPLICANT: Raunum, Laura P.W.
APPLICANT: Chung, Ming-yi
APPLICANT: Zoghbi, Huda Y.
TITLE OF INVENTION: Gene Sequence for Spinocerebellar Ataxia
Patent No. 5834183
TITLE OF INVENTION: Type I and Method for Diagnosis
NUMBER OF SEQUENCES: 85
CORRESPONDENCE ADDRESS:
ADDRESSEE: Mueeling, Raasch, Gebhardt & Schwappach, P.A.
STREET: P.O. Box 581415

```

```

CITY: Minneapolis
STATE: MN
COUNTRY: USA
ZIP: 55458-1415
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/267, 803B
FILING DATE: 28-JUN-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: McCormack, Myra H.
REGISTRATION NUMBER: 36,602
REFERENCE/DOCKET NUMBER: 110.00030120
TELECOMMUNICATION INFORMATION:
TELEPHONE: 612-305-1217
TELEFAX: 612-305-1228
INFORMATION FOR SEQ ID NO: 47:
SEQUENCE CHARACTERISTICS:
LENGTH: 66 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-267-803B-47

Query Match          0.2%; Score 38; DB 3; Length 66;
Best Local Similarity 100.0%; Pred. No. 9.1e-05;
Matches    38; Conservative   0.; Mismatches    0; Indels     0; Gaps      0

QY  7513 GTGTGTTGTGTGTGTGTGTGTGTGTATGTGTGTGTGT 7550
      |||||
Dd  29 GTGTGTGTGTGTGTGTGTGTGTGTGTATGTGTGTGTGT 66

RESULT 12
US-08-454-557C-69/C
Sequence 69, Application US/08454557C
Patent No. 5830570
GENERAL INFORMATION:
APPLICANT: de la Monte, Suzanne
APPlicant: Wands, Jack R.
TITLE OF INVENTION: Neutral Thread Protein Gene Expression and Detection
TITLE OF INVENTION: of Alzheimer's Disease
NUMBER OF SEQUENCE ADDRESSES: 121
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/454,557C
FILING DATE: 30-MAY-1995
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609.3840003
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 69:

```

SEQUENCE CHARACTERISTICS:
LENGTH: 76 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
US-08-454-557C-69

Query Match 0.2%; Score 38; DB 3; Length 76;
Best Local Similarity 100.0%; Pred. No. 8.9e-05;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12897 GGTGAACCCGTCTCTACTACAAATAATAGC 12934
|||||
DB 48 GGTGAACCCGTCTCTACTACAAATAATAGC 11

RESULT 13
US-08-340-426D-69/C

Sequence 69, Application US/08340426D
Patent No. 5948834
GENERAL INFORMATION:
APPLICANT: de la Monte, Suzanne
APPLICANT: Wands, Jack R.
TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection
TITLE OF INVENTION: Of Alzheimer's Disease
NUMBER OF SEQUENCES: 121
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/340,426D
FILING DATE: 14-NOV-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609.3840002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 69:
SEQUENCE CHARACTERISTICS:
LENGTH: 76 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
US-08-340-426D-69

Query Match 0.2%; Score 38; DB 4; Length 76;
Best Local Similarity 100.0%; Pred. No. 8.9e-05;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12897 GGTGAACCCGTCTCTACTACAAATAATAGC 12934
|||||
DB 48 GGTGAACCCGTCTCTACTACAAATAATAGC 11

RESULT 14
US-08-450-673C-69/C
Sequence 69, Application US/08450673C
Patent No. 5948888
GENERAL INFORMATION:

APPLICANT: de la Monte, Suzanne
APPLICANT: Wands, Jack R.
TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection
TITLE OF INVENTION: Of Alzheimer's Disease
NUMBER OF SEQUENCES: 121
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/450,673C
FILING DATE: 30-MAY-1995
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609.3840004
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 69:
SEQUENCE CHARACTERISTICS:
LENGTH: 76 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
US-08-450-673C-69

Query Match 0.2%; Score 38; DB 4; Length 76;
Best Local Similarity 100.0%; Pred. No. 8.9e-05;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12897 GGTGAACCCGTCTCTACTACAAATAATAGC 12934
|||||
DB 48 GGTGAACCCGTCTCTACTACAAATAATAGC 11

RESULT 15
PCT-US95-17111A-69/C

Sequence 69, Application PCT/US951711A
GENERAL INFORMATION:
APPLICANT: de la Monte, Suzanne
APPLICANT: Wands, Jack R.
TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection
TITLE OF INVENTION: Of Alzheimer's Disease
NUMBER OF SEQUENCES: 121
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/17111A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/340,426

FILING DATE: 14-NOV-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: Ludwig, Steven R.
 REGISTRATION NUMBER: 36,203
 REFERENCE/DOCKET NUMBER: 0609.3840002
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 371-2600
 TELEFAX: (202) 371-2540
 INFORMATION FOR SEQ ID NO: 69:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 76 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: both
 TOPOLOGY: both
 PCT-US93-17111A-69

Query Match 0.2%; Score 38; DB 6; Length 76;
 Best Local Similarity 100.0%; Pred. No. 8.9e-05;
 Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 12897 GGCGAAACCTGTCTCTACTAAATACAAATATAGC 12934
 ||||||||||||||||||||||||||||||||||||
 Db 48 GGCGAAACCTGTCTCTACTAAATACAAATATAGC 11

Search completed: March 29, 2000, 17:48:08
 Job time: 50350 sec



4

.

-

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 29, 2000, 06:25:57 ; Search time 5477.05 Seconds
(Without alignments)
13787.907 Million cell updates/sec

Title: US-08-852-495c-1_COPY_25000_45000
Perfect score: 20001
Sequence: 1 AGTGCAGTAGCTGCACCTC.....CAAGTTAGACATTTT 20001

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 4538634 seqs, 1887831982 residues

Word size : 0

Total number of hits satisfying chosen parameters: 130662

Minimum DB seq length: 8
Maximum DB seq length: 105

Post-processing: Listing first 45 summaries

Database : EST:
1: em_est1:*
2: em_est2:*
3: em_est3:*
4: em_est4:*
5: em_est5:*
6: em_est6:*
7: em_est7:*
8: em_est8:*
9: em_est9:*
10: em_est10:*
11: em_est11:*
12: em_est12:*
13: em_est13:*
14: em_est14:*
15: em_est15:*
16: em_est16:*
17: em_est17:*
18: em_est18:*
19: em_est19:*
20: gb_est1:*
21: gb_est2:*
22: gb_est3:*
23: gb_est4:*
24: gb_est5:*
25: gb_est6:*
26: gb_est7:*
27: gb_est8:*
28: gb_est9:*
29: gb_est10:*
30: gb_est11:*
31: gb_est12:*
32: gb_est13:*
33: gb_est14:*
34: gb_est15:*
35: gb_est16:*
36: gb_est17:*
37: gb_est18:*
38: gb_est19:*
39: gb_est20:*
40: gb_est21:*
41: gb_est22:*
42: gb_est23:*
43: gb_est24:*

44: gb_est25:*
45: gb_est26:*
46: gb_est27:*
47: gb_est28:*
48: gb_est29:*
49: gb_est30:*
50: gb_est31:*
51: gb_est32:*
52: gb_est33:*
53: gb_est34:*
54: gb_est35:*
55: gb_est36:*
56: gb_est37:*
57: gb_est38:*
58: gb_est39:*
59: gb_est40:*
60: gb_est41:*
61: gb_est42:*
62: gb_est43:*
63: gb_est44:*
64: gb_est45:*
65: gb_est46:*
66: gb_est47:*
67: gb_est48:*
68: gb_est49:*
69: gb_est50:*
70: gb_est51:*
71: gb_est52:*
72: gb_est53:*
73: gb_est54:*
74: gb_est55:*
75: gb_est56:*
76: gb_est57:*
77: gb_est58:*
78: gb_est59:*
79: gb_est60:*
80: gb_est61:*
81: gb_est62:*
82: gb_est63:*
83: gb_est64:*
84: gb_est65:*
85: gb_est66:*
86: gb_est67:*
87: gb_est68:*
88: gb_est69:*
89: gb_est70:*
90: gb_est71:*
91: gb_est72:*
92: gb_est73:*
93: gb_est74:*
94: gb_est75:*
95: gb_est76:*
96: gb_est77:*
97: gb_est78:*
98: gb_est79:*
99: gb_est80:*
100: gb_est81:*
101: gb_est82:*
102: gb_est83:*
103: gb_est84:*
104: gb_est85:*
105: gb_est86:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
c 1	51	0.3	70	25	N84707

N84707 J0579F Humu

2	47	0.2	59	AA082835	zn212i12.s
3	48	0.2	52	AA082835	zn212i12.s
4	49	0.2	72	AA191923	tr272608.x
5	44	0.2	103	AA053224	AO535224.RPCT-11-
6	43	0.2	104	AO535224	AO535224.RPCT-11-
7	40	0.2	102	BA8088	BA8088.RPCT11-4N6.
8	40	0.2	68	DD4637	DD4637.HUMSUP036
9	39	0.2	96	BD7088	BD7088.r107005.s1
10	39	0.2	100	W74144	W74144.zd0305.s1
11	38	0.2	101	AA555145	AA555145.r107004.s1
12	38	0.2	98	BD7549	BD7549.r068f10.s1
13	38	0.2	102	AA031920	AA031920.RPCT11-79
14	38	0.2	104	AA618139	AA618139.np46603.s
15	37	0.2	105	AA173240	AA173240.cm15610.y
16	37	0.2	72	AA1345123	AA1345123.td51508.x
17	37	0.2	105	AA078003	AA078003.7H12D08
18	37	0.2	105	AA832832	AA832832.at72909.x
19	36	0.2	105	AA280198	AA280198.zt04412.r
20	36	0.2	98	AA280198	AA280198.CIT-HSP-204
21	36	0.2	102	BA0126	BA0126.CIT-HSP-204
22	35	0.2	95	AA579401	AA579401.n153c01.s
23	35	0.2	96	AA747857	AA747857.nx31f09.s
24	35	0.2	99	AA026667	AA026667.z693h01.s
25	34	0.2	54	AA601314	AA601314.n015f06.s
26	34	0.2	90	AA708028	AA708028.z905601.s
27	34	0.2	90	AA208131	AA208131.qf3409.x
28	34	0.2	99	AA1583277	AA1583277.tt56607.x
29	34	0.2	101	AA235077	AA235077.z836b04.s
30	34	0.2	101	AA583697	AA583697.nm58f10.s
31	34	0.2	103	AA176421	AA176421.OVO-CT001
32	34	0.2	105	AA027613	AA027613.CITR-E1
33	33	0.2	57	AA134770	AA134770.t120b06.x
34	33	0.2	46	AA168785	AA168785.tP89003.x
35	33	0.2	78	AA0354	AA0354.yf81d05.s1
36	33	0.2	83	AA117122	AA117122.m050509.r
37	33	0.2	85	AA1254563	AA1254563.vt49801.x
38	33	0.2	85	AA1310915	AA1310915.tx76601.x
39	33	0.2	89	AA1559863	AA1559863.t032612.x
40	33	0.2	89	AA1572963	AA1572963.t032612.x
41	33	0.2	90	AA1312403	AA1312403.wm72f11.x
42	33	0.2	91	AA453508	AA453508.t123508.x
43	33	0.2	95	AA457423	AA457423.a886b02.r
44	33	0.2	95	AA578401	AA578401.n153c01.s
45	33	0.2	95	AA1202982	AA1202982.qf32610.x
46	33	0.2	99	AA486800	AA486800.rj229a06.r
47	33	0.2	100	AA564832	AA564832.n122a06.r

ALIGNMENTS

RESULT	1
LOCUS	N84707/c
DEFINITION	N84707 Human fetal heart, lambda Zap Express Homo sapiens CDNA clone J0579.5' similar to REPETITIVE ELEMENT ALU, mRNA sequence
ACCESSION	N84707
VERSION	N84707.1
KEYWORDS	EST.
SOURCE	GI:1260332
ORGANISM	human.
REFERENCE	Homo sapiens
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
TITLE	Eutheria; Primates; Catarrhini; Homnidae; Homo.
JOURNAL	l.(bases 1 to 70)
COMMENT	liew,C.C. CDNAs from fetal heart (1996) unpublished (1996) Contact: liew CC Department of Laboratory Medicine and Pathobiology University of Toronto Banting Institute, 100 College St., Toronto, Ontario, M5G1L5 Tel.: 4169788758 Fax: 4169785650 Email: liewcc@uttc.utoronto.ca Seq primer: GAAATTACCCTCAGTAAGGC.

```

FEATURES
source
location/Qualifiers
1..70
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="10579"
/clone_1lb="Human fetal heart, Lambda ZAP Express"
/lb_host="E. coli XL1-Blue"
/ncle="Vector: Lambda ZAP Express, Site_1: EcoRI, Site_2:
XhoI; mRNA was purified from human fetal hearts (8-10
weeks). cDNA was synthesized using a XhoI-Oligo dT
adaptor-primer. EcoRI adaptors were ligated, followed by
digestion with XhoI, for directional cloning into
predigested lambda ZAP Express."
BASE COUNT
24 a 18 c 15 g 13 t
ORIGIN
Query Match 0 %: Score 51; DB 25; Length 70;
Best Local Similarity 100.0%; Pred. No. 8.4e-07;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0
QY 13807 TAGAGATGGGCTTCACCATGTTAGCCAGAGATGCTCTCCATCTCCGACTT 13857
|||||
DB 57 TAGAGATGGGCTTCACCATGTTAGCCAGAGATGCTCTCCATCTCCGACTT 7

```

RESULT 2
AA082835
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
COMMENT

AA082835 59 bp mRNA
Z121912.81 Stratagene neuroepithelium NT2RAMI 937234 Homo sapiens
CDNA Clone IMAGE:548134 3' similar to contains Alu repetitive
element; mRNA sequence.
AA082835
AA082835.1 GI:1624910
EST.
human.
Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 59)
Hillier L., Lennon G., Becker M., Bonaldo M.F., Chapelli B.,
Chissole S., Dietrich N., Dubuque T., Faveiro A., Gish N.,
Hacking M., Hultman M., Kucaba T., Lacey M., Le N.,
Mardis E., Moore B., Morris M., Parsons J., Prange C., Rifkin L.,
Rowling E., Schellenberg K., Soares M.B., Tan P., Thierry-Mieg J.,
Travis J., Underwood K., Wohlmann P., Waterston R., Wilson R.
and Merris M.
Generation and analysis of 280 000 human expressed sequence tags
Genome Res. 6 (9), 807-828 (1996)
97044478

On Sep 12, 1996 this sequence version replaced gi:1404580.
Contact: Wilson RK
Washington University School of Medicine
Washington University Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LBN; contact the
IMAGE Consortium (info@image.lbl.gov) for further information.
Seq primer: -40M13 fwd. From Amer sham.
Location/Qualifiers

1..39
/organism="Homo sapiens"
/db_xref="GDB:3928650"
/db_xref="taxon:9606"
/clone="IMAGE:548134"
/clone_lib="Stratagene neuroepithelium NT2RAMI 937234"
/dev_stage="Ntera-2/Ra-M1 neuroepithelial cells"
/lab_host="SOLR (kanamycin resistant)"
/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
XhoI; Cloned unidirectionally. Primer: Oligo dT. NT2
(Ntera-2/cl.D1) precursor cells induced with RetAicof
Acid for 1 week, followed by 3 weeks in mitotic inhibitors

TITLE Venter, J.C.
JOURNAL Use of BAC End Sequences for Sequence-Ready Map Building
COMMENT Unpublished (1997) RPI11-4N6-TV
Other-GSS: RPI11-4N6-TV

Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
912 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: madams@tigr.org
Clones are derived from the human BAC library RPI1-11. For BAC
library availability, please contact Pieter de Jong
(pieter@dejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
Research Genetics (info@resgen.com). BAC end search page:
http://www.tigr.org/tldb/humgen/bac_end_search.html
Seq primer: sp6
Class: BAC ends.

FEATURES
SOURCE Location/Qualifiers

1. 102
/organism="Homo sapiens"
/db_xref="GDB:7501469"
/db_xref="taxon:9606"
/clone="RPI1-11-4N6"
/clone_11b="RPI1-11"
/sex="Male"
/cell_type="Lymphocytes"
/note="Vector PBAC3.6, Site 1: EcoRI, Site 2: EcoRI,
RPI11 Human Male BAC library"

BASE COUNT 32 a 24 c 25 g 21 t

Query Match 0.2%; Score 43; DB 81; Length 102;
Best Local Similarity 100.0%; Pred. No. 0.00033;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 12887 TGGCCACATGTTGAACCTCTCTACTATAAATACAAA 12929
Db 24 TGGCCACATGTTGAACCTCTCTACTATAAATACAAA 66

RESULT 6
LOCUS D44637/c 68 bp mRNA EST 20-PEB-1998
DEFINITION HNSUPY036 Human brain cDNA Homo sapiens cDNA clone 033-00-1, mRNA
sequence.
ACCESSION D44637
VERSION D44637.1 GI:1572112
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
Hadao, S., Ishida, Y., Tomiyasu, H., Yamamoto, K., Bates, G. P. and
Ireda, J.

REFERENCE 1 (bases 1 to 68)
TITLE Transcrip map of the human chromosome 4p16.3 consisting of 627
JOURNAL DNA Res. 3 (4), 239-255 (1996)
MEDLINE 97101646
COMMENT On May 5, 1995 this sequence version replaced g1:798369.
Contact: Shinji Hadao
Japan Science and Technology Corporation, Neurogenes Project, ICOMP
Univ. of Tokai School of Med.
Bohsedai, Isehara, Kanagawa 259-1193, Japan
Tel: 81-463-91-5095
Fax: 81-463-91-4993
Email: shinji@ng.med.u-tokai.ac.jp.

FEATURES
SOURCE Location/Qualifiers
1. 68
/organism="Homo sapiens"

/db_xref="taxon:9606"
/clone="033-00-1"
/clone_11b="Human brain cDNA"
/tissue_type="brain"

BASE COUNT 14 a 16 c 25 g 13 t

Query Match 0.2%; Score 40; DB 28; Length 68;
Best Local Similarity 100.0%; Pred. No. 0.0042;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 8495 GCAATTCCTGCTCAGCCTCCGAGTAGCTGGGATTAC 8534
Db 68 GCAATTCCTGCTCAGCCTCCGAGTAGCTGGGATTAC 29

RESULT 7
LOCUS R67088 96 bp mRNA EST 30-MAY-1995
DEFINITION Y130H05.s1 Soares placenta NB2HP Homo sapiens cDNA clone
IMAGE:140793 3', mRNA sequence.
ACCESSION R67088
VERSION R67088.1 GI:839726
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
Haller, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,
Holman, M., Huilman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,
Pearson, J., RIKLIN, L., Kohling, T., Soares, M., Tan, F.,
Trevaskis, E., Waterston, R., Williamson, A., Wohlmann, P. and
Wilson, R.

The WashU-Merck EST Project
Unpublished (1995)
On Apr 5, 1995 this sequence version replaced g1:760768.
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Insert Size: 1093
High quality sequence stops: 65
Source: IMAGE Consortium, LNLN.
This clone is available royalty-free through LNLN; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
Insert length: 1093 Std Error: 0.00
Seq primer: Promega -21m13
High quality sequence stop: 65.

FEATURES
SOURCE Location/Qualifiers

1. 96
/organism="Homo sapiens"
/db_xref="GDB:349405"
/db_xref="taxon:9606"
/clone="IMAGE:140793"
/clone_11b="Soares placenta NB2HP"
/sex="Female"

/dev_stage="placenta obtained at birth (full term)"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: Placenta; Vector: p773D (pharmacia) with a
modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer (5'
AAGTGGAGATTGCGCGCCGAGATTATTTTATTTT 3'),
double-stranded cDNA was ligated to Eco RI adaptors
(pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified p773 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M. Patricia Bonaldo.

BASE COUNT 25 a 24 c 21 g 25 t 1 others

AUTHORS

Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chiapella, B., Chisoe, S., Dietrich, N., Duboulet, T., Favello, A., Gish, M., Hawkins, M., Hultman, M., Kucaba, T., Lacy, M., Le, N., Mardis, E., Moore, B., Morris, M., Parsons, J., Prange, C., Ritzkin, L., Roeding, T., Schellenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J., Trevisan, E., Underwood, K., Wohlman, P., Waterston, R., Wilson, R., and Maria, M.

TITLE

Generation and analysis of 280,000 human expressed sequence tags
JOURNAL MEDLINE
Genome Res. 6 (9), 807-828 (1996)
97044478

FEATURES

On Oct 18, 1995 this sequence version replaced g1:1026289.
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: estewatson.wustl.edu
Insert Size: 174
Source: IMAGE Consortium, LNL
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
Seq primer: Promea-21m13
Location/Qualifiers

1..98
/organism="Homo sapiens"
/db_xref="GDB:3864371"
/db_xref="taxon:9606"
/clone="IMAGE:238987"
/clone_lib="Weizmann Olfactory Epithelium"
/sex="Female"
/tissue_type="Olfactory epithelium"
/divergence="35 year old"
/lab_host="SOLR cells (kanamycin resistant)"
/note="Organ: nose; Vector: pBluescript SK-; Site_1:
ECORI; Site_2: XhoI; Cloned unidirectionally. Primer:
Oligo dt. Olfactory epithelium, normal. Average insert
size: 0.8 kb; Uni-ZAP XR Vector. Library constructed by N.
Walker, D. Lancet, Weizmann Institute of Science. -5'
adaptor sequence: 5' GAAATCGCGCAGG 3' -3' adaptor
sequence: 5' CTCGAGTTTCTTTTCTTTT 3'."
BASE COUNT 19 a 39 c 23 g 17 t
ORIGIN

Query Match 0.2%; Score 38; DB 24; Length 98;
Best Local Similarity 100.0%; Pred. No. 0.016;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8481 CCTCCAGTCAAGCAATCTCTGCTCCAGCCGCC 8518
|||||
DB 50 CTTCCAGGTTCAAGCAATCTCTGCTCCAGCCGCC 87

RESULT 11
A0319270 102 bp DNA GSS 06-MAY-1999
LOCUS A0319270
DEFINITION RPII11-98B22.TV RPII-11 Homo sapiens genomic clone RPII-11-98B22,
genomic survey sequence.
ACCESSION A0319270
VERSION A0319270
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 102)
Adams, M.D., Rounsley, S.D., Zhao, S., Base, S., Liner, K., Golden, K.,
Berry, K., Granger, D., Suh, E., Wible, C., de Jong, P., and Venter, J.C.
Unpublished (1998)
Other-GSSs: RPII11-98B22.TV
Contact: Shaying Zhao, William Nierman, Mark Adams
Department of Eukaryotic Genomics

The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208
Email: hbe@igf.org
Clones are derived from the human BAC library RPII-11. For BAC
library availability, please contact Pieter de Jong
(pieter@dejong.med.buitalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buitalo.edu/ordering) or from
Research Genetics (info@resgen.com). BAC end search page:
http://www.igf.org/tdb/hungen/bac_end_search.html
Seq primer: Sp6
Class: BAC ends.
Location/Qualifiers

1..102
/organism="Homo sapiens"
/db_xref="GDB:7537293"
/db_xref="taxon:9606"
/clone="RPII-11-98B22"
/clone_lib="RPII-11"
/sex="Male"
/cell_type="T lymphocytes"
/note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
RPII11 Human Male BAC library"
BASE COUNT 26 a 28 c 31 g 17 t
ORIGIN

Query Match 0.2%; Score 38; DB 100; Length 102;
Best Local Similarity 100.0%; Pred. No. 0.016;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 17275 CCACCGGCTGGGCTCCCAAGTGTGGATTACAG 17312
|||||
DB 91 CCACCGGCTGGGCTCCCAAGTGTGGATTACAG 54

RESULT 12
AA614379 104 bp mRNA EST 06-OCT-1997
LOCUS AA614379
DEFINITION np46c03.81 NCI-CGAP Brl.1 Homo sapiens cDNA clone IMAGE:1129348 3'
similar to contains Alu repetitive element; mRNA sequence.
ACCESSION AA614379
VERSION AA614379.1 GI:2465575
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 104)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
On May 5, 1995 this sequence version replaced g1:798056.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
www-bio.lnl.gov/bdnp/image/image.html

Seq primer: -40m13 fwd. ET from Amersham
High quality sequence step 52.
Location/Qualifiers
1..104
/organism="Homo sapiens"
/db_xref="taxon:9606"
FEATURES
source

	/clone="IMAGE:1129348"	
	/clone_lib="NCI_GCAP_Brl.1"	
	/sex="Female, pooled"	
	/tissue_type="breast"	
	/lab_host="DH10B"	
	(note="Vector: pRT7D-Pac (Pharmacia) with a modified polylinker; 1st strand cDNA was prepared from pooled bulk breast tumor tissue, and was then primed with a Not I - oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pRT73 vector. Library is not normalized. (The normalized version of this library is NCI CGAP Br-2.) Library was constructed by Bento Soares and M. Fatima Bonaldo."	
BASE COUNT	19 a	26 c 29 g 30 t
ORIGIN		
Query Match	0.2%; Score 38;	DB 36; Length 104;
Best Local Similarity	100.0%;	Pred. No. 0.016;
Matches 38;	Conservative 0;	Mismatches 0;
	Indels 0;	Gaps 0;
Oy 8635 CCGCCTCCGGCCTCCCAAGTGCTGGGATTACAGGC GTG 8672 		
Dd 67 CCGCCTCCGGCCTCCCAAGTGCTGGGATTACAGGC GTG 104		
RESULT 13		
A1752407/c		
LOCUS	A1752407	105 bp mRNA EST 22-JUN-1999
DEFINITION	cni5d10.v1 Normal Human Trabecular Bone Cells Homo sapiens CDNA	
ACCESSION	A1752407	
VERSION	A1752407.1 GI:5130671	
KEYWORDS	EST.	
SOURCE	human.	
ORGANISM	Homo sapiens	
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo. 1 (bases 1 to 105) Jia,L.B., Young,M.F., Touchman,J.W., Bouffard,G.G., Beckstrom-Sternberg,S.M., Green,E.D., Powell,J.I., Yang,L.M., Robey,P.G., Hotchkiss,R.N. and Franccomano,C.A. SCAP: The Skeletal Genome Anatomy Project Unpublished (1997) On Dec 20, 1995 this sequence version replaced gi:1130959. Contact: Libin Jia Medical Genetics Branch National Human Genome Research Institute 1010ClO1, 9000 Rockville Pike, Bethesda, MD 20892-1267, USA Tel.: 301-402-4877 Fax: 301-496-7157 Email: libin@helix.nih.gov DNA sequencing and analyses by National Institutes of Health Intramural Sequencing Center (NISC). Plate: 15 row: d column: 10 Seq primer: M13RP1 reverse primer (ABI).	
TITLE	JOURNAL	
COMMENT		
FEATURES	Location/Qualifiers	
source	1..105	
	/organism="Homo sapiens"	
	/db_xref="taxon:9606"	
	/clone_NHTBC_cni5d10"	
	/clone_lib="Normal Human Trabecular Bone Cells"	
	/sex="female"	
	/tissue_type="Bone"	
	/cell_type="Trabecular Bone Cells"	
	/lab_host="SURF"	
	/note="Organ: Hip; Vector: pBluescript; Site_1: ECORI"	
BASE COUNT	34 a	22 c 26 g 22 t 1 others
ORIGIN		
Query Match	0.2%; Score 38;	DB 51; Length 105;

Best Local Similarity	100.0%	Pred. No. 0.016%	Matches 38	Conservative 0	Mismatches 0	Indels 0	Gaps 0
Oy	8646	TCCCAAAGTCTGGGATTACAGGCGTGAACACACACG	8663				
Db	105	TCCCAAAGTCTGGGATTACAGGCGTGAACACACACG	68				
RESULT	14						
LOCUS	AI1349123/c						
DEFINITION	AI1349123	72 bp	mRNA	EST	30-DEC-1998		
ACCESSION	ta51908.x2	NCI-CGAP_Lu25	Homo sapiens	CDNA clone	IMAGE:2047646	3',	
VERSION	AI1349123						
KEYWORDS	AI1349123.1	GI:4086329					
SOURCE	EST.						
ORGANISM	human.						
REFERENCE	Homo sapiens						
AUTHORS	Eukaryota: Metazoa; Chordata; Craniata; Vertebrata; Mammalia;						
TITLE	Eutheria; Primates; Catarrhini; Hominoidea; Homo.						
COMMENT	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap .						
JOURNAL	National Cancer Institute, Cancer Genome Anatomy Project (CGAP),						
CONTACT	Tumor Gene Index						
TEL	Unpublished (1997)						
DB	Contact: Robert Strausberg, Ph.D.						
FEATURES	Tel: (301) 496-1530						
SOURCE	Email: Robert.Strausberg@nih.gov						
Seq primer: -40UP from GIBCO.	CDNA Library Preparation: David B. Krizman, Ph.D.						
Location/Qualifiers	NCI-CGAP Arrayed by: I.M.A.G.E. Consortium, LLNL						
1..72	CDNA sequencing by: Washington University Genome Sequencing Center						
/organism="Homo sapiens"	clone distribution: NCI-CGAP clone distribution information can be						
/db_xref="taxon:9606"	found through the I.M.A.G.E. Consortium/LLNL at:						
/clone="IMAGE:2047646"	www-bio.llnl.gov/dbfp/image/image.html						
/clone_lib="NCI CGAP_Lu25"							
/tissue_type="bronchioalveolar carcinoma"							
/dev_stage="adult"							
/lab_host="DH10B"							
/note="Organ: Lung; Vector: PAMPI; mRNA made from lung							
carcinoma tissue, cDNA made by oligo-dT priming.							
directionally cloned. Size-selected on agarose gel,							
average insert size 500 bp. Primary library,							
non-amplified.							
BASE COUNT	40 a	8 c	17 g	7 t			
ORIGIN							
Query Match		0.2%	Score 37;	DB 45;	Length 72;		
Best Local Similarity	100.0%;	Pred. No. 0.042;					
Matches 37;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;			
Oy	18764	TTTTTTTTTTTTTTTTTTTTTTCAGACAGAGTCTC	18800				
Db	62	TTTTTTTTTTTTTTTTTTTTTTTTCAGACAGAGTCTC	26				
RESULT	15						
LOCUS	AA078003	105 bp	mRNA	EST	24-SEP-1999		
DEFINITION	7H12D08	Chromosome 7	HeLa	CDNA Library	Homo sapiens	CDNA clone	
ACCESSION	AA078003						
VERSION	AA078003.1	GI:1837477					
KEYWORDS	EST.						
SOURCE	human.						
ORGANISM	Homo sapiens						

REFERENCE

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

AUTHORS

1 (bases 1 to 105)
Touchman, J.W., Bouffard, G.G., Weintraub, L.A., Idol, J.R., Wang, L.,
Robbins, C.M., Nussbaum, J.C., Lovett, M., and Green, E.D.

TITLE

2006 expressed-sequence tags derived from human chromosome 7-enriched cDNA libraries

JOURNAL

Genome Res. 7 (3), 281-292 (1997)

MEDLINE

97228905

COMMENT

On Apr 14, 1993 this sequence version replaced gi:693433.

Contact: Eric D. Green

Genome Technology Branch

National Human Genome Research Institute/NIH

49 Convent Dr., MSC4431, Building 49, Room 2A08, Bethesda, MD 20892

Tel.: 3014020201

Fax: 3014024735

Email: egreen@nhgri.nih.gov

Plate: 12 row: D column: 08

Seq primer: -21M13 (ABI).

Location/Qualifiers

1..105

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="7H12D08"

/sex="female"

/cell_line="Hela cell line; ATCC"

/lab_host="E. coli strain DH5 alpha"

/note="Vector: PAMP10; cDNA was generated from cytoplasmic RNA using a mixture of random DNA hexamers and oligo(dT). From this pool of cDNA, human chromosome 7-enriched cDNA was isolated by direct cDNA selection using chromosome 7 genomic DNA (cosmids). The resulting direct-selected cDNA was cloned into a plasmid vector using a non-directional uracil DNA glycosylase (UDG)-mediated cloning strategy."

20 a 33 c 23 g 29 t

BASE COUNT
ORIGIN

Query Match 0.28; Score 37; DB 28; Length 105;
Best Local Similarity 100.0%; Pred. No. 0.034; 0; Indels 0; Gaps 0;
Matches 37; Conservative 0; Mismatches 0;

QY

10241 AATTGCTGACCCAGAGCAGAGGTTGCAGTGAGC 10277

DB

37 AATTGCTGACCCAGAGCAGAGGTTGCAGTGAGC 1

Search completed: March 29, 2000, 14:49:02
Job time: 41596 sec

KW Synthetic gene; synthesis; polyester; L-amino acid analogue; clips;
 KW stop codon; stop signal; monomer unit; lactate; glycolate; staples;
 KW translation initiation; modified; Met-tRNA^f; cyanogen bromide; pins;
 KW co-polymer; sutures; drug delivery device; screws; ds.
 OS Synthetic.
 PN CA2112716-A.
 PD 01-JUL-1994.
 PF 31-DEC-1993; 112716.
 PR 31-DEC-1992; US-999520.
 PR (US) US SURGICAL CORP.
 PI Gruskin EA;
 DR MPI: 94-272324/34.
 PR New biocompatible medical devices and articles - using a lactic
 acid-glycolic acid polyester in which monomer units are
 PT individually and specifically determined
 PS Example 2: Page 53: 67pp: English.
 CC The sequences given in Q71814-16 represent fragments of the
 CC synthetic polyester genes given in Q71801-02. These fragments were
 CC used in the production of the full length polyester gene in the plasmid
 CC pMAL-72. This plasmid was used so that the polyester could be expressed
 CC in an E. coli cell free medium, as pMAL-p2 is recognised by E. coli
 CC ribozymes and contains the necessary initiator sequences and the
 CC template for the polypeptide leader. The resulting plasmid pMAL-2c
 CC contains the complete polyester coding sequence ligated in frame with
 CC the leader sequence. The expressed sequences represent synthetic
 CC genes which were used in the synthesis of polyesters derived from the
 CC L-amino acid analogues of an amino acid except Pro. Of the three stop
 CC codons, UAA, UAG and UGA, one of these stop codons is reversed for the
 CC stop signal for the polyester. The remaining two stop codons are
 CC available for encoding the monomer units of the polyester. Lactate
 CC and glycolate are encoded by UAA and UAG, respectively. Initiation of
 CC translation always starts at an AUG codon, therefore the first lactate
 CC of the chain may be encoded by an AUG codon. This is accomplished
 CC through the use of a specially modified Met-tRNA^f modified to carry
 CC lactate instead of Met. Alternatively, a Met is incorporated in the
 CC first position of the polymer chain. The resulting polymer is treated
 CC with cyanogen bromide to remove the Met. Co-polymers produced in this
 CC manner may be used to produce sutures, staples, clips, drug delivery
 CC devices, pins and screws.
 CC Sequence 81 BP: 47 A; 6 C; 4 G; 24 T;

[illegible]

01-JUL-1994.
31-DEC-1993: 112716.
31-DEC-1992: US-9995520.
(USSU) US SURGICAL CORP.
GruSkin EA;
WPI: 94-272324/34.
New biocompatible medical devices and articles - using a lactic acid-glycolic acid polyester in which monomer units are individually and specifically determined
Example 1: Page 45; 67pp: English.
This sequence represents a fragment of the plasmid pSPORT 1c which contains the 5' fragment of the synthetic genes which encode polyesters, given in Q71801-02, under the control of the E. coli trpA transcription initiator and followed by PstI and HindIII restriction sites. A further fragment of the synthetic genes were ligated into pSPORT 1c to give plasmid pSPORT 1d (see also Q71810-11), with the final fragment being inserted in a third round of ligation to give pSPORT 1e (see also Q71812-13). The E. coli trpA initiator sequence was used to express the full length sequences given in Q71801-02 in an E. coli cell free medium. The expressed sequences represent synthetic genes which were used in the synthesis of polyesters derived from the α -amino acid analogues of an amino acid except Pro. Of the three stop codons, UAA, UAG and UGA, one of these stop codons is reversed for the stop signal for the polyester. The remaining two stop codons are available for encoding the monomer units of the polyester. Lactate and glycolate are encoded by UAA and UAG, respectively. Initiation of translation always starts at an AUG codon, therefore the first lactate of the chain may be encoded by an AUG codon. This is accomplished through the use of a specially modified Met-tRNA^f modified to carry lactate instead of Met. Alternatively, a Met is incorporated in the first position of the polymer chain. The resulting polymer is treated with cyanogen bromide to remove the Met. Co-polymers produced in this manner may be used to produce sutures, staples, clips, drug delivery devices, pins and screws.
Sequence 103 BP; 58 A; 7 C; 9 G; 29 T;

```

Query Match Similarity      0.1%; Score 41; DB 1; Length 103;
Best Local Similarity      100.0%; Pred. No. 0.0035;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 27146 AATATATATATATATATATATATATATATATATATATATATATATATAA 27186
|||||
|||||
Db 23 AATATATATATATATATATATATATATATATATATATATATATATATAA 63

RESULT 7
033621
AC 033621 standard; DNA; 53 BP.
DT 02-FEB-1993 (first entry)
DE Microsatellite sequence from clone GBFSR.
KW PCR; selection; primers; OPPIRIM; breeding; cattle; parentage;
KM genetic mapping; traits; amplification; ss.
OS Bos taurus.
PN W09213102-A.
PD 06-AUG-1992.
PF 15-JAN-1992; U00340.
PR 15-JAN-1991; US-642342.
PA (GENM-) GENMARK.
PI Georges M. Massey JM;
PT WPI; 92-284684/34.
PR Polymorphic bovine DNA markers - used in genetic identification,
PT gene mapping, and selective breeding
PT Table 7; Page 178; 51pp; English.
CC The sequence is that of a bovine microsatellite sequence obcd. by
CC screening a library of bovine MboI DNA fragments of between
CC 250 and 500 bp with an (AC)15 and a (TC)15 oligonucleotide probe.
CC One out of 50 clones cross-hybridised. Assuming independent
CC distribution of microsatellites and MboI sites, the frequency of
CC (T6)n >9 microsatellites in the bovine genome is estimated at >100,
CC 000. The sequence information for ca. 230 such bovine microsatellites
CC is summarised in the specification and indexed herein (see below).
CC The sequences upstream and downstream of the microsatellite sequence

```


Mon Apr 3 08:24:26 2000

us-08-852-495c-1_copy_40000_70000.rng

Page 7



2
3
4

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 29, 2000, 17:48:08 ; Search time 340.37 Seconds
(without alignments)
10552.676 Million cell updates/sec

Title: US-08-852-495C-1_COPY_40000_70000

Perfect score: 30001
Sequence: 1 CTCGACACGAGTTCAGCTA.....AACACAGTTGAGATTTGA 30001

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 214294 seqs, 59861574 residues

Word size : 0

Total number of hits satisfying chosen parameters: 349160

Minimum DB seq length: 8
Maximum DB seq length: 105

Post-processing: Listing first 45 summaries

Database :

Issued_Patents_NA: *
1: /cgn2_6/prodata/1/ina/5A_COMB.seq: *
2: /cgn2_6/prodata/1/ina/5B_COMB.seq: *
3: /cgn2_6/prodata/1/ina/5C_COMB.seq: *
4: /cgn2_6/prodata/1/ina/5D_COMB.seq: *
5: /cgn2_6/prodata/1/ina/6_COMB.seq: *
6: /cgn2_6/prodata/1/ina/PCUT09_COMB.seq: *
7: /cgn2_6/prodata/1/ina/backfile1.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	39	0.1	39	2	US-08-632-673B-1
2	39	0.1	50	2	US-08-241-465B-11
3	38	0.1	68	3	US-08-417-174-42
4	38	0.1	58	3	US-08-231-65A-40
5	38	0.1	58	4	US-09-007-961-40
6	38	0.1	70	4	US-08-776-944-12
7	38	0.1	75	4	US-08-776-944-13
8	38	0.1	80	2	US-07-920-281C-25
9	37	0.1	37	1	US-08-113-646A-44
10	37	0.1	40	4	US-08-771-624B-1
11	37	0.1	40	4	US-08-440-209-4
12	37	0.1	50	1	US-08-420-443-1
13	37	0.1	55	1	US-08-113-646A-41
14	37	0.1	69	2	US-08-702-344-7
15	37	0.1	69	2	US-08-702-344-22
16	37	0.1	84	2	US-08-664-596B-3
17	37	0.1	84	2	US-08-738-367-3
18	37	0.1	90	2	US-08-677-944-1
19	37	0.1	90	2	US-08-677-944-2
20	37	0.1	98	4	US-08-088-658-42
21	37	0.1	98	4	US-08-471-907A-42
22	37	0.1	105	3	US-08-735-381-2
23	35	0.1	85	1	US-08-120-827-97
24	35	0.1	85	1	US-08-478-675-97
25	34	0.1	84	3	US-08-454-557C-91
26	34	0.1	84	3	US-08-454-557C-91

27	34	0.1	84	4	US-08-340-426D-91	Sequence 91, Appl
28	34	0.1	84	4	US-08-340-426D-91	Sequence 91, Appl
29	34	0.1	84	4	US-08-450-673C-91	Sequence 91, Appl
30	34	0.1	84	4	US-08-450-673C-91	Sequence 91, Appl
31	34	0.1	84	6	PCT-US95-17111A-91	Sequence 91, Appl
32	34	0.1	84	6	PCT-US95-17111A-91	Sequence 91, Appl
33	33	0.1	85	3	US-08-454-557C-92	Sequence 92, Appl
34	33	0.1	85	4	US-08-340-426D-92	Sequence 92, Appl
35	33	0.1	85	4	US-08-450-673C-92	Sequence 92, Appl
36	33	0.1	85	6	PCT-US95-17111A-92	Sequence 92, Appl
37	32	0.1	47	4	US-08-778-494B-114	Sequence 114, Appl
38	32	0.1	50	1	US-08-381-572-20	Sequence 20, Appl
39	32	0.1	50	2	US-08-592-820-20	Sequence 20, Appl
40	32	0.1	51	4	US-08-582-562A-8	Sequence 8, Appl
41	32	0.1	51	4	US-08-778-494B-8	Sequence 8, Appl
42	32	0.1	51	4	US-08-859-998-1373	Sequence 1373, Ap
43	32	0.1	52	4	US-08-778-494B-111	Sequence 111, Appl
44	32	0.1	55	4	US-08-582-562A-16	Sequence 16, Appl
45	32	0.1	55	4	US-08-778-494B-16	Sequence 16, Appl

ALIGNMENTS

RESULT 1
US-08-632-673B-1/C
Sequence 1, Application US/08632673B
Patent No. 5712098
GENERAL INFORMATION:
APPLICANT: Tsuchihashi, Zenta
APPLICANT: Gnlrke, Andreas
APPLICANT: Thomas, Winston J.
APPLICANT: Dryna, Dennis T.
APPLICANT: Ruddy, David
APPLICANT: Wolff, Roger K.
APPLICANT: Feder, John N.
TITLE OF INVENTION: HEREDITARY HEMOCHROMATOSIS
TITLE OF INVENTION: DIAGNOSTIC MARKERS AND DIAGNOSTIC METHODS
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/632,673B
FILING DATE: 16-APR-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Smith, William M.
REGISTRATION NUMBER: 30,223
REFERENCE/DOCKET NUMBER: 01957-000410
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 326-2400
TELEFAX: (415) 326-2422
INFORMATION FOR SEQ. ID NO. 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 39 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-632-673B-1


```

SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/231.565A
FILING DATE: 22-APR-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: CAROL M. GUPTI
REGISTRATION NUMBER: 37,341
REFERENCE/DOCKET NUMBER: 2026-41240
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 756-4800
TELEFAX: (212) 751-6849
TELEX: 421792
INFORMATION FOR SEQ ID NO: 40:
SEQUENCE CHARACTERISTICS:
LENGTH: 58
TYPE: NUCLEOTIDE
STRANDEDNESS: DOUBLE
TOPOLOGY: UNKNOWN
MOLECULE TYPE: CDNA

```

```
Query Match      0.1% Score 38; DB 3; length 58;
Best Local Similarity 100.0%; Pred. No. 0.00066;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

1 RESULT 5
2 US-09-007-961-40
3 ; Sequence 40, Application US/09007961
4 ; Patent No. 5994523
5 ;
6 ; GENERAL INFORMATION:
7 ; APPLICANT: KAWAKAMI, YUTAKA; ROSENBERG,
8 ; APPLICANT: STEVEN A.
9 ; TITLE OF INVENTION: MELANOMA ANTIGENS AND
10 ; TITLE OF INVENTION: THEIR USE IN DIAGNOSTIC AND THERAPEUTIC
11 ; TITLE OF INVENTION: METHODS
12 ; NUMBER OF SEQUENCES: 43
13 ;
14 ; CORRESPONDENCE ADDRESS:
15 ; ADDRESSEE: MORGAN & FINNEGAN
16 ; STREET: 345 PARK AVENUE
17 ; CITY: NEW YORK
18 ; STATE: NEW YORK
19 ;
20 ; COUNTRY: USA
21 ;
22 ; ZIP: 10154
23 ;
24 ; COMPUTER READABLE FORM:
25 ; MEDIUM TYPE: FLOPPY DISK
26 ; COMPUTER: IBM PC COMPATIBLE
27 ; OPERATING SYSTEM: PC-DOS/MS-DOS
28 ; SOFTWARE: ASCII
29 ;
30 ; CURRENT APPLICATION DATA:
31 ; APPLICATION NUMBER: US/09/007,961
32 ; FILING DATE:
33 ; CLASSIFICATION:
34 ;
35 ; PRIOR APPLICATION DATA:
36 ; APPLICATION NUMBER: 08/231,565
37 ; FILING DATE: 22-APR-1994
38 ; ATTORNEY/AGENT INFORMATION:
39 ; NAME: CAROL M. GRUPPI
40 ; REGISTRATION NUMBER: 37,341
41 ;
42 ; REFERENCE/DOCKET NUMBER: 2026-4124
43 ;
44 ; TELECOMMUNICATION INFORMATION:
45 ; TELEPHONE: (212) 758-4800
46 ; TELEFAX: (212) 751-6849
47 ; TELEX: 421792
48 ;
49 ; INFORMATION FOR SEQ ID NO: 40:
50 ; SEQUENCE CHARACTERISTICS:
51 ; LENGTH: 58

```

```

TYPE: NUCLEOTIDE
STRANDEDNESS: DOUBLE
TOPOLOGY: UNKNOWN
MOLECULE TYPE: CDNA
US-09-007-961-40

```

```

Query Match      0.1%; Score 38; DB 4; Length 58;
Best Local Similarity 100.0%; Pred. No. 0.00066;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 13600 CTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 13637
      |||
Db 18 CTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 55

```

```

      RESULT        6
US-08-776-944-12
; Sequence 12, Application US/08776944
; Patent No. 5923520
; GENERAL INFORMATION:
; APPLICANT: TULLY, GILLIAN
; APPLICANT: SULLIVAN, KEVIN
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCING METHOD
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHAYE P.C.
; STREET: 1100 NORTH GLEEBOE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/776,944
; FILING DATE: 14-FEB-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9417211.1
; FILING DATE: 25-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: CRAMFORD, ARTHUR R.
; REGISTRATION NUMBER: 25,327
; REFERENCE/DOCKET NUMBER: 1498-92
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4006
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 70 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; ORGANELLE: Mitochondrion
US-08-776-944-12

Query Match          0.1%; Score 38; DB 4; Length 70;
Best Local Similarity 100.0%; Pred. No. 0.00064;
Matches   38; Conservative    0; Mismatches    0; Indels    0; Gaps    0;

Q# 13601 TTTTGGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTG 13638
|||||
TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTG 41

```



```

: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/113,646A
: FILING DATE: 31-AUG-1993
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/084,406
: FILING DATE: 10-AUG-1987
: ATTORNEY/AGENT INFORMATION:
: NAME: WILSON, MARY J.
: REGISTRATION NUMBER: 32,955
: REFERENCE/DOCKET NUMBER: 1579-20
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (703) 816-4000
: TELEFAX: (703) 816-4100
: TELEX: 200797 NIXN UR
: INFORMATION FOR SEQ ID NO: 44:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 37 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: RNA (genomic)
: US-08-113-646A-44

```

Query Match	0.1%	Score 37;	DB 1;	length 37;
Best Local Similarity	100.0%	Pred. No. 0.0015;		
Matches	37;	Conservative	0;	Mismatches 0; Indels 0; Gaps 0;
QY	13602	TTTTTTTTTTTTTTTTTTTTTTTTTTTGG	13638	
Db	37	TTTTTTTTTTTTTTTTTTTTTTTTTTTGG	1	

RESULT 10
 US-08-771-624B-1
 : Sequence 1, Application US/08771624B
 : Patent No. 5914230
 : GENERAL INFORMATION:
 : APPLICANT: Liu, Yen Ping
 : APPLICANT: Patel, Rajesh D.
 : APPLICANT: Kurn, Nurith
 : APPLICANT: Lin, Claire
 : APPLICANT: Rose, Samuel J.
 : APPLICANT: Ullman, Edwin F.
 : TITLE OF INVENTION: Homogeneous Amplification and Detection
 : of Nucleic Acids
 : NUMBER OF SEQUENCES: 24
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Behring Diagnostics GmbH, c/o Dade Behring Inc.
 : STREET: 1717 Deerfield Road
 : CITY: Deerfield
 : STATE: Illinois
 : COUNTRY: US
 : ZIP: 60015-0778
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: Patentin Release #1.0, Version #1.30
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/08/771.624B
 : FILING DATE: 20-DEC-1996
 : CLASSIFICATION: 435
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: US 60/009,090
 : FILING DATE: 22-DEC-1995
 : CLASSIFICATION: 435
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Ruzsala, Lois K.
 : REGISTRATION NUMBER: 39,074
 : REFERENCE/DOCKET NUMBER: 1030
 : TELECOMMUNICATION INFORMATION:
 :

TELEPHONE: (847) 267-5366
TELEFAX: (847) 267-6024
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 40 base pairs
TYPE: nucleic acid
SPANDENESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomid)
HYPOTHETICAL: NO
ANTI-SENSE: NO
PACKAGE TYPE: C-terminal
US-08-771-624B-1

Query Match	0.1%	Score 37	DB 4	Length 40
Best Local Similarity	100.0%	Pred. No.	0.0015	
Matches 37	Conservative 0	Mismatches 0	Indels 0	Gaps 0

Qy 13601 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 13637
 |||||
Ddb 1 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 37

```

US-08-440-209-4
: Sequence 4, Application US/08440209
: Patent No. 5922857
: GENERAL INFORMATION:
:   APPLICANT: Han, Jang H
:   APPLICANT: Spaete, Richard R
:   TITLE OF INVENTION: Methods and Compositions for Controlling
:   TITLE OF INVENTION: Translation of HCV Proteins
:   NUMBER OF SEQUENCES: 5
:   CORRESPONDENCE ADDRESS:
:     ADDRESSEE: Wolf, Greenfield, and Sacks P.C.
:     STREET: 600 Atlantic Avenue
:     CITY: Boston
:     STATE: MA
:     COUNTRY: USA
:     ZIP: 02210
: COMPUTER READABLE FORM:
:   MEDIUM TYPE: Floppy disk
:   COMPUTER: IBM PC compatible
:   OPERATING SYSTEM: PC-DOS/MS-DOS
:   SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
:   APPLICATION NUMBER: US/08/440,209
:   FILING DATE: 12-MAY-1995
:   CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
:   APPLICATION NUMBER: 08/477,895
:   FILING DATE: 06-JUN-1995
:   APPLICATION NUMBER: US 08/128,583
:   FILING DATE: 28-SEP-1993
: ATTORNEY/AGENT INFORMATION:
:   NAME: Janluk, Anthony J
:   REGISTRATION NUMBER: 29,809
:   REFERENCE/DOCKET NUMBER: C0772/7004
: TELECOMMUNICATION INFORMATION:
:   TELEPHONE: 617-720-3500
:   TELEFAX: 617-720-2441
: INFORMATION FOR SEQ ID NO: 4:
:   SEQUENCE CHARACTERISTICS:
:     LENGTH: 40 base pairs
:     TYPE: nucleic acid
:     STRANDEDNESS: single
:     TOPOLOGY: linear
:     MOLECULE TYPE: cDNA
:     HYPOTHEICAL: NO
:     ANTI-SENSE: YES
US-08-440-209-4

```



```

SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/702,344
FILING DATE:
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Brown, Scott A.
REGISTRATION NUMBER: 32,724
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 498-8224
TELEFAX: (617) 876-5851
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 69 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-702-344-7

```

[illegible]

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

MOLECULE TYPE: CDNA
; US-08-702-344-22

	Query Match	0.1%	Score 37:	DB 2:	Length 69:
	Best Local Similarity	100.0%	Pred.	No. 0.0014:	
Matches	37,	Conservative	0;	Mismatches	0; Indels 0; Gaps 0;
Cy	13601 TTTTTCCTTTTTTTTTTTTTTTTTTTTTTTT	13637			
Dd	69 TTTTTCCTTTTTTTTTTTTTTTTTTTTTTTT	33			

Search completed: March 30, 2000, 00:47:59
Job time: 75541 sec

Mon Apr 3 08:24:26 2000

us-08-852-495c-1_copy_40000_70000.rni

Page 8

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 29, 2000, 14:49:02 ; Search time 8179.5 Seconds
(without alignments)
13848.486 Million cell updates/sec

Title: US-08-852-495c-1_COPY_40000_70000

Sequence: 1 GTGTACATCAGCTTCAGCTA.....AACACACTTGCAATTGA 30001

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 4538634 seqs, 1887831982 residues

Word size : 0

Total number of hits satisfying chosen parameters: 130662

Minimum DB seq length: 8

Maximum DB seq length: 105

Post-processing: Listing first 45 summaries

Database : EST:
1: em_est1:*
2: em_est2:*
3: em_est3:*
4: em_est4:*
5: em_est5:*
6: em_est6:*
7: em_est7:*
8: em_est8:*
9: em_est9:*
10: em_est10:*
11: em_est11:*
12: em_est12:*
13: em_est13:*
14: em_est14:*
15: em_est15:*
16: em_est16:*
17: em_est17:*
18: em_est18:*
19: em_est19:*
20: gb_est1:*
21: gb_est2:*
22: gb_est3:*
23: gb_est4:*
24: gb_est5:*
25: gb_est6:*
26: gb_est7:*
27: gb_est8:*
28: gb_est9:*
29: gb_est10:*
30: gb_est11:*
31: gb_est12:*
32: gb_est13:*
33: gb_est14:*
34: gb_est15:*
35: gb_est16:*
36: gb_est17:*
37: gb_est18:*
38: gb_est19:*
39: gb_est20:*
40: gb_est21:*
41: gb_est22:*
42: gb_est23:*
43: gb_est24:*

44: gb_est25:*
45: gb_est26:*
46: gb_est27:*
47: gb_est28:*
48: gb_est29:*
49: gb_est30:*
50: gb_est31:*
51: gb_est32:*
52: em_est20:*
53: em_est21:*
54: em_est22:*
55: em_est23:*
56: em_est24:*
57: em_est25:*
58: em_est26:*
59: gb_est33:*
60: gb_est34:*
61: gb_est35:*
62: gb_est36:*
63: gb_est37:*
64: gb_est38:*
65: em_est27:*
66: em_est28:*
67: em_est29:*
68: em_est30:*
69: gb_est39:*
70: gb_est40:*
71: gb_est41:*
72: gb_est42:*
73: gb_est43:*
74: gb_est44:*
75: em_est31:*
76: em_est32:*
77: em_est33:*
78: em_est34:*
79: gb_est45:*
80: gb_est46:*
81: gb_est47:*
82: gb_est48:*
83: em_est35:*
84: em_est36:*
85: em_est37:*
86: em_est38:*
87: gb_est49:*
88: gb_est50:*
89: gb_est51:*
90: gb_est52:*
91: gb_est53:*
92: em_est39:*
93: em_est40:*
94: em_est41:*
95: em_est42:*
96: em_est43:*
97: em_est44:*
98: em_est45:*
99: gb_est54:*
100: gb_est55:*
101: em_est46:*
102: gb_est56:*
103: gb_est57:*
104: gb_est58:*
105: gb_est59:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
c 1	58	0.2	65	81	B36140 HS-1038-A1-

C	2	53	0.2	91	38	AA780764	ac68f12.	AA780764	ac68f12.
C	3	51	0.2	92	40	AA916300	on22c08.s	AA916300	on22c08.s
C	4	47	0.2	72	62	AT914923	tr27e08.x	AT914923	tr27e08.x
C	5	44	0.1	95	24	H61099	xy51c12..r1	H61099	xy51c12..r1
C	6	43	0.1	68	28	AA083205	zn26h12.x	AA083205	zn26h12.x
C	7	43	0.1	68	28	AA083205	zn26h12.x	AA083205	zn26h12.x
C	8	43	0.1	93	44	AA038446	AA038446	AA038446	AA038446
C	9	43	0.1	101	25	NS4156	NS4156	NS4156	NS4156
C	10	42	0.1	101	25	NS4156	NS4156	NS4156	NS4156
C	11	42	0.1	95	24	H61099	xy51c12..r1	H61099	xy51c12..r1
C	12	41	0.1	73	51	AA072757	AA072757	AA072757	AA072757
C	13	41	0.1	87	20	D030989	H0MGS01971	D030989	H0MGS01971
C	14	41	0.1	98	24	H67549	xy68f10..s1	H67549	xy68f10..s1
C	15	41	0.1	102	22	H60869	xy08c10..s1	H60869	xy08c10..s1
C	16	41	0.1	102	22	H60869	xy08c10..s1	H60869	xy08c10..s1
C	17	40	0.1	72	42	AT653742	AT653742	AT653742	AT653742
C	18	40	0.1	82	39	AA841232	xy36k02.x	AA841232	xy36k02.x
C	19	40	0.1	82	39	AA841232	xy36k02.x	AA841232	xy36k02.x
C	20	40	0.1	97	74	AA188576	AA188576	AA188576	AA188576
C	21	40	0.1	155	100	AAQ282107	AAQ282107	AAQ282107	AAQ282107
C	22	39	0.1	58	35	C35772	C35772	C35772	C35772
C	23	39	0.1	49	49	AA047660	AA047660	AA047660	AA047660
C	24	39	0.1	61	39	AA824000	V93J609..t	AA824000	V93J609..t
C	25	39	0.1	61	68	HSM002661	HSM002661	HSM002661	HSM002661
C	26	39	0.1	71	49	AT638201	AT638201	AT638201	AT638201
C	27	39	0.1	71	50	F37855	F37855	F37855	F37855
C	28	39	0.1	72	40	AA927569	AA927569	AA927569	AA927569
C	29	39	0.1	77	64	AA083168	AA083168	AA083168	AA083168
C	30	39	0.1	79	44	AT298440	AT298440	AT298440	AT298440
C	31	39	0.1	79	44	AT298440	AT298440	AT298440	AT298440
C	32	39	0.1	82	45	F34490	F34490	F34490	F34490
C	33	39	0.1	82	45	F34490	F34490	F34490	F34490
C	34	39	0.1	84	49	AT349814	AT349814	AT349814	AT349814
C	35	39	0.1	84	49	AT349814	AT349814	AT349814	AT349814
C	36	39	0.1	93	33	AA428622	AA428622	AA428622	AA428622
C	37	39	0.1	94	62	AT915297	AT915297	AT915297	AT915297
C	38	39	0.1	96	30	R28974	R28974	R28974	R28974
C	39	39	0.1	96	30	AA199893	AA199893	AA199893	AA199893
C	40	39	0.1	100	26	W74144	W74144	W74144	W74144
C	41	39	0.1	101	35	AA555145	AA555145	AA555145	AA555145
C	42	39	0.1	101	46	AT436429	AT436429	AT436429	AT436429
C	43	39	0.1	102	64	AA087020	AA087020	AA087020	AA087020
C	44	39	0.1	102	64	AA087020	AA087020	AA087020	AA087020
C	45	39	0.1	102	68	HSM010460	HSM010460	HSM010460	HSM010460

ALIGNMENTS

RESULT	1
B36140.AC	
LOCUS	
DEFINITION	B36140 65 bp DNA GSS 17-OCT-1997
ACCESSION	HS-1038-A1-D06-MR.ab1 C1T Human Genomic Sperm Library C Homo sapiens genomic clone Plate-CT 820 Col-11 Row-G, genomic survey sequence.
VERSION	B36140
KEYWORDS	B36140.1 GI:2535509
SOURCE	GSS.
ORGANISM	human.
REFERENCE	Homo sapiens
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
TITLE	1 (bases 1 to 65) Mallatras/G.G., Zackrone,K.D., Smith,T., Tipton,S., Schmidt,S., Traisoff,R., Abajian,C., Blanchard,A., West,A. and Hood,L.E. Construction of a Characterized Clone Resource for Genomic Sequencing: Generation and Preliminary Analysis of 20,000 Sequence Tagged Connectors Unpublished (1997)
JOURNAL COMMENT	Contact: Mallatras GG, zackrone KD, Hood L University of Washington Seattle, WA 98195, USA Tel.: (206) 616-8744

```

FEATURES
source
    1. .65
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /clone_plate="CT 820 Col-11 Row-G"
    /clone_lib="CTR Human Genomic Sperm Library C"
    /sex="M"
    /note="Organ: sperm; Vector: pBelOBAcl1; BAC Clones in
    E-coli DH10b"
BASE COUNT      25 a      18 c      10 g      12 t
ORIGIN
Query Match
Best Local Similarity 100.0%; Pred. No. 5.4e-08;
Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps
0.2%; Score 58; DB 81; Length 65;
CY 9842 GGTATATTTTGTATTTTGTAGAGACGGGATTCACCATGTTGGTCACGCTGCTC 9899
DB 59 GGTTAATTTTGTATTTTGTATTTGTAGACGGGATTCACCATGTTGGTCACGCTGCTC 2

```

```

RESULT      2
LOCUS       AA780764/c
DEFINITION  AA780764          91 bp      mRNA                      EST           05-FEB-1998
acbf8f12.s1 Strataegus fetal retina g37202 Homo sapiens CDNA clone
IMAGE:8677791 3' similar to contains Alu repetitive element/, mRNA
sequence.

ACCESSION   AA780764
VERSION     AA780764
KEYWORDS    EST
SOURCE      human.
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE   1 (bases 1 to 91)
AUTHORS     Hillier,L., Allen,M., Bowles,L., Dubuque,T., Giesel,G., Jost,S.,
            Kruman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M.,
            Martin,J., Moore,B., Schellenberg,R., Stepien,M., Tan,F.,
            Thelasing,B., White,Y., Wyllie,T., Waterston,R. and Wilson,R.
TITLE       WashU-NCI human EST project
JOURNAL     Unpublished (1997)
COMMENT     On Sep 12, 1996 this sequence version replaced gi:1407381.
CONTACT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel.: 314 286 1800
            Fax: 314 286 1810
            Email: serwatson.wustl.edu
            This clone is available royalty-free through LINT : contact the
            IMAGE Consortium (info@image.lind.gov) for further information.
            Seq primer: -40ml3 fwd. Et from Amersham.
            Location/Qualifiers
                1..91
                    /organism="Homo sapiens"
                    /db_xref="taxon:9606"
                    /clone_image="IMAGE:867791"
                    /clone_lib="Stratagene fetal retina 937202"
                    /sex="mixed"
                    /lab_host="SOLR (kanamycin resistant)"
                    /note="Vector: pBluescript SK-. Site1: EcoRI; Site2:
XhoI; Cloned unidirectionally. Primer: Oligo dT. Pooled
retinal tissue. Average insert size: 1.0 kb. Mini-TAP XR
Vector: ~5' adaptor sequence: 5' GAATTCGCGCAGGAG 3' ~3'
adaptor sequence: 5' CTCGACTTTTTTTTTTTTTTTT 3'"
FEATURES             source

```

Query Match	0.24;	Score 53;	DB 38;	Length 91;
Best Local Similarity	100.0%;	Pred. No. 1,7e+06;		
Matches 53; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	27059	CGGGAGGGCGAGGTTCGACAGTGCAGCCGAGATTTCGCGACACGACACTCCAGGCTGG	27121	
Db	74	CGGGAGGGCGAGGTTCGACAGTGCAGCCGAGATTTCGCGACACGACACTCCAGGCTGG	22	

RESULT	3
AA916300/c	
LOCUS	AA916300 92 bp mRNA EST 17-JUN-1998
DEFINITION	similar2c08.s1 NCI-CGAP_lus Homo sapiens cDNA clone IMAGE:1557422 3
ACCESSION	AA916300
VERSION	AA916300.1 GT:3055692
KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens

REFERENCE	1 (bases 1 to 92)
AUTHORS	NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap .
TITLE	National Cancer Institute, Cancer Genome Anatomy Project (CGAP)
JOURNAL	Tumor Gene Index
COMMENT	Unpublished (1997)
	On May 5, 1995 this sequence version replaced gi:797742.

Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/ULNU at: www-bio.llnl.gov/dbfp/image/image.html

```

Insert Length: 1519      Std Error: 0.00
Seq primer: -40m13 fwd.  ET from Amersham
High quality sequence stop: 61.
      Location/Qualifiers
          1. 92
FEATURES
    source

```

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1557422"
/clone_1lb="NCI_CGAP_Lu5"
/tissue_type="carcinoid"
/lab_host="DH10B"
/note="Organ: Lung; Vector: pT7T3D-Pac (Pharmacia) with a
modified polylinker; 1st strand cDNA was prepared from
neuroendocrine lung carcinoid, and was then primed with a
Not I - oligo(dT) primer. Double-stranded cDNA was ligated
to Eco RI adaptors (Pharmacia), digested with Not I and
cloned into the Not I and Eco RI sites of the modified
pT7T3 vector. Library is normalized. Library was
constructed by Bento Soares and M. Fatima Bonaldo. "
15 a 34 c 21 g 22 t

```

```

Query Match      0.2%; Score 51; DB 40; Length 92;
Best Local Similarity 100.0%; Pred. No. 7e-06;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

0Y 11995 CAGGAGGCTGAGGCAGGAGAAATTCCTTGAACCCGGGAGGTGAGGTTCAG 12045
      |||||
Db 85 CAGGAGGCTGAGGCAGGAGAAATTCCTTGAACCCGGGAGGTGAGGTTCAG 35

```

[illegible]

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

1 (bases 1 to 72)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
On May 1, 1997 this sequence version replaced gi:205935.

Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/dbfp/image/image.html

```

FEATURES
  source
    Location/Qualifiers
      1..72
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /clone="IMAGE:2219558"
        /clone_11b="NCL CGAP OV23"
        /tissue_type="tumor, 5 pooled (see description)"
        /lab_host="DH10B"
        /note="Organ: ovary; Vector: pCMV-SPORT6; Site_1: SalI;
        Site_2: NotI; Cloned unidirectionally. Primer: oligo dT
        Average insert size 1.35 kb. Tumor types include: mixed
        Mullerian tumor, papillary serous, clear cell, spindle
        cell. All are primary tumors, metastasis positive. Life
        Technologies catalog #: 11534-013"
BASE COUNT
  8 a
  11 c
  15 g
  38 t

```

	Query Match	0.2%	Score 47:	DB 62:	Length 72:	
	Best Local Similarity	100.0%	Pred. No.	0 00014:		
	Matches 47:	Conservative 0:	Mismatches 0:	Indels 0:	Gaps 0:	
QY	3759	TTTTTTT	TTTTTTT	TGACAGACAGATCCT	CACCT	3805
D6	4	TTTTTTT	TTTTTTT	TGACAGACAGATCCT	CACCT	50

RESULT	5
H61099/c	
LOCUS	
DEFINITION	H61099 95 bp mRNA EST 06-OCT-1995 y5512.t1 Soares fetal liver spleen INFLS Homo sapiens cdna clone
ACCESSION	y5512.t1
VERSION	H61099
KEYWORDS	IMAGE:208822 5', mRNA sequence.
SOURCE	H61099.1 GI:1013931
ORGANISM	EST.
REFERENCE	human.
AUTHORS	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Euarcharia; Primates; Catarrhini; Hominoidea; Homo. 1 (bases 1 to 95) Hillier,L., Clark,N., Dubague,T., Elliston,K., Hawkins,M., Holman,R., Hultman,M., Kucaba,T., Le,M., Lennon,G., Maize,M.,

TITLE
 The Mashu-Merck EST Project
 Unpublished (1995)
 On Sep 21, 1992 this sequence version replaced g1:279332.

High quality sequence stops: 87
Source: IMAGE Consortium, LNL
This clone is available royalty-free through LNL ; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
Insert Length: 1755 Std Error: 0.00
Seq primer: M13RPI
High quality sequence stop: 87.

FEATURES

source

Location/Qualifiers
1. .95

/organism="Homo sapiens"
 /db_xref="GDB:3777953"
 /db_xref="taxon:9606"
 /clone IMAGE:208822"
 /clone_lib="Soares fetal liver spleen INFs"
 /sex="male"
 /dev_stage="30 week post conception fetus"
 /lab_host="DH10B (ampicillin resistant)"
 /note="Origin: Liver and Spleen. Vector: p773D (Pharmacia)
 with a modified polylinker. Site.1: Pac I. Site.2: Eco RI.
 1st strand cDNA was primed with a Pac I oligo (d1)
 3' AACTGGAGATTAATTAATGAATCTTTTCTTTTCTTTT 3',
 double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Pac I and cloned into the Pac I
 and Eco RI sites of the modified p773 vector. Library
 went through one round of normalization. Library
 constructed by Bento Soares and M.Fátima Bonaldo."

```
Query Match Score 44; DB 24; Length 95;
Best Local Similarity 100.0%; Pred. No. 0.0011;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 20223 TATATATATATATATATATATATATATATATATG 20266
db 95 TATATATATATATATATATATATATATATATATATG 52
```

[illegible]

REFERENCE
AUTHORS

TITLE Generation and analysis of 280,000 human expressed sequence tags

JOURNAL	Genome Res. 6 (9), 807-828 (1996)
MEDLINE	97044478
COMMENT	On Sep 12, 1996 this sequence version replaced gi:1397700.

Email: estevenson.west@du.edu
WARNING: There is evidence that suggests that the 384-well parent plate of this clone contains both human and mouse derived clones. Thus, the origin of this clone is uncertain. This caution should be kept in mind should you use this clone.

This clone is available royalty-free through LMD ; contact the IMAGE Consortium (info@image.lmd.jussieu.fr) for further information.
Seq primer : -28M13 rev2 from Amersham
High quality sequence stop : 54.

FEATURES
source

Location/Qualifiers
1. .68

```

/organism="Homo sapiens"
/db_xref="GDB:3927131"
/db_xref="taxon:9606"
/clone_image="348615"
/clone_id="Stratagene neuroepithelium NT2RAM1 937234"
/dev_stage="Neura-2/R4-M1 neuroepithelial cells"
/lab_host="SODR (Kanamycin resistant)"
/notes="vector: plasmidscript SK-; Site_1: EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer: Oligo dr. NT2 (Neura-2/c1.D1) precursor cells induced with Retinoic Acid for 1 week, followed by 3 weeks in mitotic inhibitors (Repaglate #2). Average insert size: 1.1 kb; anti-zap XR vector; -5' adaptor sequence: 5' GAATTCGACGACG 3' -3' adaptor sequence: 5' CTCGAGTCTTTTTTTTTTTTTTTT 3'"

```

[illegible]

RESULT	7	
LOCUS	AA082205/c	
DEFINITION	AA082205 68 bp mRNA	EST 23-DEC-1997
ACCESSION	z20h11.r1	Stratagene neuroepithelium NT2.RMI
VERSION	CDNA clone IMAGE:548615 5', mRNA sequence.	937234 Homo sapiens
KEYWORDS	AA082205.1	GI:1624264
SOURCE	EST.	
ORGANISM	Homo sapiens	human.

REFERENCE
AUTHORS

REFERENCE	AUTHORS
1 (bases 1 to 68)	
	Hillier, L., Lennon, G., Becker, M., Donald, M. F., Chlapelli, B.,
	Dietch, N., Dubage, T., Favella, A., Gish, N.,
	Chiscoe, S., Hultman, M., Kucaba, T., Lacey, M., Le, M.,
	Hawkins, M., Moore, B., Morris, M., Parsons, J., Prings, C., Rifkin, L.,
	Morris, E., Scellenberg, K., Soares, M. B., Tan, F., Thierry-Mieg, J.,
	Rohlfing, T., Schellenberg, K., Soares, M. B., Tan, F., Thierry-Mieg, J.,
	Trevaskis, E., Underwood, K., Wohldmann, P., Waterston, R., Wilson,
	and Marra, M.

TITLE Generation and analysis of 280,000 human expressed sequence tags
 JOURNAL Genome Res. 6 (9), 807-828 (1996)
 MEDLINE 97044478
 COMMENT On Sep 12, 1996 this sequence version replaced gi:1397700.

Washington University School of Medicine

	RESULT	12	AU072757	73 bp	mRNA	EST	24-JUN-1999
	LOCUS	AU072757					
	DEFINITION	AU072757 Dictyostellium discoideum SS (H.Urushihara) Dictyostellium discoideum cDNA clone SSA765; mRNA sequence.					
	ACCESSION	AU072757					
	VERSION	AU072757.1	GI:5179178				
	KEYWORDS	EST.					
	SOURCE	Dictyostelium discoideum.					
	ORGANISM	Dictyostelium discoideum.					
	REFERENCE	Eukaryota; Dictyostellidae; Dictyostellium.					
	TITLE	1 (bases 1 to 73)					
	AUTHORS	Urushihara,H.					
	JOURNAL	Developmental cDNA in Dictyostellium discoideum (1999)					
	COMMENT	Unpublished (1999) On Jun 5, 1998 this sequence version replaced gi::3188495. Contact: Hideko Urushihara Institute of Biological Sciences University of Tsukuba 3-3-10 Ten-nodai, Tsukuba, Ibaraki 305, Japan Email: d402hu@esakura.cc.tsukuba.ac.jp PROJECT - Dictyostelium discoideum cDNA project in Japan.					
	FEATURES	source					
		1..73	/organism="Dictyostelium discoideum"				
			/strain="AX4"				
			/db_xref="taxon:44689"				
			/clone_id="SSA765"				
			/dev_stage="slug"				
	BASE COUNT	46 a	0 c	1 g	26 t		
	ORIGIN						
	Query Match	0.1%;	Score 41;	DB 51;	Length 73;		
	Best Local Similarity	100.0%;	Pred. No. 0.011;				
	Matches 41:	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;		
Oy	27146	AAAAAAAAATTAATAATAATAATAATAATAATAATAATAAA	27186				
Dd	15	AAAAATTAATAATAATAATAATAATAATAATAATAATAA	55				
	RESULT	13	D20989	87 bp	mRNA	EST	30-JUL-1996
	LOCUS	D20989/c					
	DEFINITION	HMGSO1971 Human promyelocyte Homo sapiens CDNA clone mp0383 3', mRNA sequence.					
	ACCESSION	D20989					
	VERSION	D20989.1	GI:504809				
	KEYWORDS	EST.					
	SOURCE	human.					
	ORGANISM	Homo sapiens					
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homindae; Homo.					
	REFERENCE	1 (bases 1 to 87)					
	AUTHORS	Okubo,K., Fukushima,A., Yoshii,J., Niyama,T., Kojima,Y., Yoshinari,H., Arimoto,J. and Matsubara,K.					
	TITLE	Gene expression of human promyelocytic cell line HL60 before and after induction of differentiation. A new application of 3'directed cDNA sequencing (1993)					
	JOURNAL	Contact: Okubo,K., Fukushima,A., Yoshii,J., Niyama,T., Kojima,Y., Yoshinari,H., Arimoto,J. and Matsubara,K.					
	COMMENT	Institute for Molecular and Cellular Biology Osaka University 3-1 Yamada-Oka,Suita, Osaka 565,Japan. Location/Qualifiers 1..87 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="mp0383"					
	BASE COUNT	19 a	39 c	23 g	17 t		
	ORIGIN						
	note="Female."						
	promyelocyte."						
	Base Count	13 a	23 c	17 g	34 t		
	ORIGIN						
	Query Match	0.1%;	Score 41;	DB 20;	Length 87;		
	Best Local Similarity	100.0%;	Pred. No. 0.0099;				
	Matches 41:	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;		
Oy	20914	GTGGCTCACACTGTTAATCCAGCACTTGGAGCGTGAGC	20954				
Dd	51	GTGGCTCACACTGTTAATCCAGCACTTGGAGCGTGAGC	11				
	RESULT	14	H67549	98 bp	mRNA	EST	11-DEC-1995
	LOCUS	H67549					
	DEFINITION	yv68f10.s1 Weizmann Olfactory Epithelium Homo sapiens cDNA clone IMAGE:238987 3' similar to contains Alu repetitive element;contains MER22 repetitive element ; , mRNA sequence.					
	ACCESSION	H67549					
	VERSION	H67549.1	GI:1114431				
	KEYWORDS	EST.					
	SOURCE	human.					
	ORGANISM	Homo sapiens					
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homindae; Homo.					
	REFERENCE	1 (bases 1 to 98)					
	AUTHORS	Hallier,L., Lennon,G., Becker,M., Bernaldo,M.F., Chiappelli,B., Chissoe,S., Dietrich,N., Dubouque,T., Favallo,A., Gish,W., Hawkins,M., Hultman,M., Kucaba,I., Lacy,M., Le,M., Le,N., Marlis,E., Moore,R., Morris,M., Parsons,J., Prange,C., Rifkin,L., Ronfling,T., Schellenberg,K., Soares,W.B., Tan,F., Thierry-Veg,J., Trevaaskis,E., Underwood,K., Wohlmann,P., Waterston,R., Wilson.R. and Marra,M.					
	TITLE	Generation and analysis of 280,000 human expressed sequence tags Genome Res. 6 (9), 807-828 (1996)					
	JOURNAL MEDLINE	On Oct 18, 1995 this sequence version replaced gi:1026289.					
	COMMENT	Contact: Wilson RK Washington Universlty School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel.: 314 286 1800 Fax: 314 286 1810 Email: estewatson.wustl.edu Insert Size: 174 Source: IMAGE Consortium, LLNL This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (infoimage.llnl.gov) for further information. Seq primer: Promega -21m13. Location/Qualifiers 1..98 /organism="Homo sapiens" /db_xref="GBS:3864371" /db_xref="taxon:9606" /clone="IMAGE:238987" /clone_id="Weizmann Olfactory Epithelium" /sex="Female" /tissue_type="olfactory epithelium" /dev_stage="35 year old" /lab_host="SOLR cells (kanamycin resistant)" /note="Organ					

Mon Apr 3 08:24:27 2000

us-08-852-495c-1_copy_40000_70000.rst

Page 8

Query Match 0.1%; Score 41; DB 24; Length 98;
Best Local Similarity 100.0%; Pred. No. 0.0092;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 9768 CCTCCAGGTTCAAGCAATCTCGCTCAGCCGCCAG 9808
50 CCTCCAGGTTCAAGCAATCTCGCTCAGCCGCCAG 90

Db 50 ATATATATATATATATATATATATATATATATATATA 90
Search completed: March 29, 2000, 20:38:58
Job time: 62592 sec

RESULT 15

R60869

LOCUS

DEFINITION

IMAGE:42595 3', mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

On Apr 18, 1995 this sequence version replaced g1:775532.

Contact: Wilson RK

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: estewatson.wustl.edu

Insert Size: 1405

High quality sequence stops: 76 Source: IMAGE Consortium, LNL, This

clone is available royalty-free through LNL; contact the IMAGE

Consortium (info@image.llnl.gov) for further information.

Insert Length: 1405 Std Error: 0.00

Seq primer: SP6

High quality sequence stop: 76.

Location/Qualifiers

1..102

/organism="Homo sapiens"

/db_xref="GB:415186"

/db_xref="taxon:9606"

/clone_image:42595"

/clone_lib="Soares Infant brain INIB"

/sex="female"

/dev_stage="73 days post natal"

/lab_host="DH10B (ampicillin resistant)"

/note="Organ: whole brain; Vector: lacMid BA; Site: 1; Not

I: Site: 2; Hind III; 1st strand cDNA was primed with a Not

I - oligo(dT) primer [5'

ACTGAGAGATTCGCGCCGACGAGATTTTCTTTTCTTTT 3'];

double-stranded cDNA was ligated to Hind III adaptors

(Pharmacia), digested with Not I and directionally cloned

into the Not I and Hind III sites of the lacMid BA vector.

Library went through one round of normalization. Library

constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 35 a 9 c 5 g 52 t 1 others

ORIGIN

Query Match 0.1%; Score 41; DB 22; Length 102;

Best Local Similarity 100.0%; Pred. No. 0.0091;

Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20224 ATATATATATATATATATATATATATATATATA 20264

|||||

Mon Apr 3 08:24:27 2000

us-08-852-495c-1_copy_40000_70000.rst

Page 9



•

•

•

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 30, 2000, 00:22:34 ; Search time 13413.1 Seconds

(Without alignments)
-6791.482 Million cell updates/sec

Title: US-08-852-495c-1_COPY_65000_95000

Perfect score: 30001
Sequence: 1 TTATATATGACAAATTAAT.....TACAGTACTCTGCAGTTT 30001Scoring table: OLIGO_NUC
Gapop 60.0 , Gapept 60.0

Searched: 821193 seqs, -1518192014 residues

Word size : 0

Total number of hits satisfying chosen parameters: 345848

Minimum DB seq length: 8
Maximum DB seq length: 105

Post-processing: Listing first 45 summaries

Database : GenEmbl.*
1: gb_ba1:*
2: gb_ba2:*
3: gb_om:*
4: gb_ov:*
5: gb_pat:*
6: gb_ph:*
7: gb_pl1:*
8: gb_pl2:*
9: gb_pl3:*
10: gb_pl4:*
11: gb_pl5:*
12: gb_pl6:*
13: gb_pl7:*
14: gb_pl8:*
15: gb_pl9:*
16: gb_pl10:*
17: gb_pl11:*
18: gb_pl12:*
19: gb_pl13:*
20: gb_pl14:*
21: gb_pl15:*
22: gb_pl16:*
23: gb_pl17:*
24: gb_pl18:*
25: gb_pl19:*
26: gb_pl20:*
27: gb_pl21:*
28: gb_pl22:*
29: gb_pl23:*
30: gb_pl24:*
31: gb_pl25:*
32: gb_pl26:*
33: gb_pl27:*
34: gb_pl28:*
35: gb_pl29:*
36: gb_pl30:*
37: gb_pl31:*
38: gb_pl32:*
39: gb_pl33:*
40: gb_pl34:*
41: gb_pl35:*
42: gb_pl36:*
43: gb_pl37:*

44: gb_pl38:*
45: gb_pl39:*
46: gb_pl40:*
47: gb_pl41:*
48: gb_pl42:*
49: gb_pl43:*
50: gb_pl44:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	58	0.2	76	5 AR051499	AR051499 Sequence
2	55	0.2	60	5 AR051487	AR051487 Sequence
3	48	0.2	48	9 HUMH116	D28457 Human mRNA
4	41	0.1	104	9 HUMALCE272	M87889 Human carcl
5	36	0.1	57	13 HUMUT1543A	L30907 Human STS U
6	36	0.1	66	40 AF087511	AF087511 Homo sapi
7	35	0.1	57	13 HUMUT1543A	L30907 Human STS U
8	35	0.1	85	5 AR051522	AR051522 Sequence
9	34	0.1	42	5 HUMALUNNC	L36838 Homo sapien
10	34	0.1	42	5 AR051521	AR051521 Sequence
11	33	0.1	35	5 A25212	A25212 Inter-Alu s
12	33	0.1	35	24 E09140	E09140 Synthetic D
13	33	0.1	40	5 A68621	A68621 Sequence 1
14	33	0.1	65	9 HUMALUNNCB	L36836 Homo sapien
15	33	0.1	69	13 HUMUT578A	L39139 Human STS U
16	33	0.1	85	5 AR051522	AR051522 Sequence
17	32	0.1	51	10 S62605	S62605 Cl-Inhibito
18	32	0.1	63	3 DOGSR7SL1	M30845 Dog signal
19	32	0.1	67	9 HUMALUNCD	L36843 Homo sapien
20	32	0.1	80	9 HUMBRKFC	M36133 Human alpha
21	31	0.1	66	10 S54071	S54071 hypoxanthin
22	31	0.1	67	13 HUMUT44A	L29843 Human STS U
23	31	0.1	90	9 HUMLDLRL	K03535 Human low d
24	31	0.1	90	9 HUMLDLRL	K03535 Human low d
25	31	0.1	90	9 HUMLDLRL	M15365 Human low d
26	31	0.1	100	11 HSU67848	U67848 Human beta-
27	31	0.1	101	13 HUMUT832A	L31225 Human STS U
28	31	0.1	104	5 A08899	A08899 H.sapiens (
29	30	0.1	40	5 A68621	A68621 Sequence 1
30	30	0.1	66	40 AF087511	AF087511 Homo sapl
31	30	0.1	67	13 HUMUT44A	L29843 Human STS U
32	30	0.1	73	5 A08915	A08915 H.sapiens (
33	30	0.1	80	9 HUMBRKFC	M36134 Human alpha
34	30	0.1	80	9 HUMBRKFC	M36135 Human alpha
35	30	0.1	91	13 G38859	G38859 TA50 Plasm
36	30	0.1	95	4 GGR10H7	X78616 G.gallus ge
37	30	0.1	99	13 HUMUT7692A	L30306 Human STS U
38	30	0.1	101	13 HUMUT832A	L31225 Human STS U
39	30	0.1	102	5 A08911	A08911 H.sapiens (
40	30	0.1	103	9 HUMALCE221	M87886 Human carcl
41	30	0.1	105	13 G32655	G32655 A009L30 Hum
42	30	0.1	105	13 HS015XB3	Z50878 H.sapiens (
43	29	0.1	60	5 AR051490	AR051490 Sequence
44	29	0.1	63	13 HUMUT5302A	L30829 Human STS U
45	29	0.1	65	9 HUMALUNNCB	L36836 Homo sapien

ALIGNMENTS

RESULT	1				
LOCUS	AR051499		76 bp	DNA	
DEFINITION	Sequence	69	from patent US 5830670.		
ACCESSION	AR051499				
VERSION	AR051499.1		GI:5974863		
				PAT	29-SEP-1999

KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 76)
AUTHORS de la Monte, S., and Wands, J.R.
TITLE Neural thread protein gene expression and detection of Alzheimer's disease
JOURNAL Patent: US 5830670-A 69 03-NOV-1998;
FEATURES Location/Qualifiers
source 1..76
/organism="unknown"
BASE COUNT 15 a 19 c 18 g 24 t
ORIGIN

Query Match 0.2%; Score 58; DB 5; Length 76;
Best Local Similarity 100.0%; Pred. No. 7.8e-19;
Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 27685 GCCCGCTAATTTTGTATTTAGTAGACAGAGGTTTCACCGTTGGCCAGATG 27742
DB 6 GCCCGCTAATTTTGTATTTAGTAGACAGAGGTTTCACCGTTGGCCAGATG 63

RESULT 2
LOCUS AR051487 60 bp DNA PAT 29-SEP-1999
DEFINITION Sequence 57 from patent US 5830670.
ACCESSION AR051487
VERSION AR051487.1 GI:5974851
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 60)
AUTHORS de la Monte, S., and Wands, J.R.
TITLE Neural thread protein gene expression and detection of Alzheimer's disease
JOURNAL Patent: US 5830670-A 57 03-NOV-1998;
FEATURES Location/Qualifiers
source 1..60
/organism="unknown"
BASE COUNT 12 a 14 c 15 g 19 t
ORIGIN

Query Match 0.2%; Score 55; DB 5; Length 60;
Best Local Similarity 100.0%; Pred. No. 2.9e-17;
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 27685 GCCCGCTAATTTTGTATTTAGTAGACAGAGGTTTCACCGTTGGCCAGG 27739
DB 6 GCCCGCTAATTTTGTATTTAGTAGACAGAGGTTTCACCGTTGGCCAGG 60

RESULT 3
LOCUS HUMH116 48 bp mRNA PRI 07-FEB-1999
DEFINITION Human mRNA for histone H1D, 5'UTR (sequence from the 5'cap to the start codon).
ACCESSION D28457
VERSION D28457.1 GI:461197
KEYWORDS histone H1D.
SOURCE Homo sapiens (Library:HT-1080/pKAI) fibrosarcoma cell-line HT-1080
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 48)
AUTHORS Kato, S.
TITLE Direct Submission
JOURNAL Submitted (03-FEB-1994) to the DDBJ/EMBL/GenBank databases. Setshl

Kato, Sagami Chemical Research Center, Genetic Engineering Section;
4-4-1 Nishi-Onuma, Sagamihara, Kanagawa 229, Japan
(E-mail: btrn00121@biotech.net.com, Tel: 0427-42-4791 (ex. 415),
Fax: 0427-49-7631)
REFERENCE 2 (bases 1 to 48)
AUTHORS Kato, S.
JOURNAL Unpublished (1994)
REFERENCE 3 (sites)
AUTHORS Sekine, S., Oh, S.W., Kim, N.S., Umezawa, Y., Abe, N.,
Yokoyama, Kobayashi, M., and Aoki, T.
TITLE Construction of a human full-length cDNA bank
JOURNAL Gene 150 (2), 243-250 (1994)
MEDLINE 95121910
COMMENT Submitted (03-Feb-1994) to DDBJ by:
Setshi Kato
Genetic Engineering Section
Sagami Chemical Research Center
4-4-1 Nishi-Onuma
Sagamihara, Kanagawa 229
Japan
Phone: 0427-42-4791
Fax: 0427-49-7631

FEATURES
source
1..48
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_line="HT-1080"
/clone_1id="HT-1080/pKAI"
/tissue_type="fibrosarcoma"
1..42
43..>48
/product="histone H1D"
/protein_id="BA05823.1"
/db_xref="GI:433162"
/translation="MS"
BASE COUNT 8 a 14 c 9 g 17 t
ORIGIN

Query Match 0.2%; Score 48; DB 9; Length 48;
Best Local Similarity 100.0%; Pred. No. 1.4e-13;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12760 ATCGGCGCTTGGCAGCTTGTAACCGAGTTTGTGATCTCAACATGTC 12807
DB 1 ATCGGCGCTTGGCAGCTTGTAACCGAGTTTGTGATCTCAACATGTC 48

RESULT 4
LOCUS HUMALCE272/C 104 bp ss-RNA PRI 15-APR-1994
DEFINITION Human carcinoma cell-derived Alu RNA transcript, clone CE272.
ACCESSION M87889
VERSION M87889.1 GI:174875
KEYWORDS Alu repeat.
SOURCE Homo sapiens male embryo carcinoma cDNA to other RNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
REFERENCE 1 (bases 1 to 104)
AUTHORS Sinnett, D., Richer, C., Deragon, J.-M., and Labuda, D.
TITLE Alu RNA transcripts in human embryonal carcinoma cells. Model of post-transcriptional selection of master sequences
JOURNAL J. Mol. Biol. (1992) In press
FEATURES Location/Qualifiers
source 1..104
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_line="M87889"
/dev_stage="embryo"
/sex="male"
/tissue_type="carcinoma"

BASE COUNT 22 a 26 c 37 g 19 t
ORIGIN

Query Match 0.1%; Score 41; DB 9; Length 104;
Best Local Similarity 100.0%; Pred. No. 7.3e-07;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2681 CCTCAGCCTCCCAAGTGTGGATTACAGAGTGCAC 2721
DB 56 CCTCAGCCTCCCAAGTGTGGATTACAGAGTGCAC 16

RESULT 5
HUMUT1543A/C 57 bp DNA SRS 17-JAN-1995
LOCUS HUMUT1543A 57 bp DNA SRS 17-JAN-1995
DEFINITION Human SRS UT1543, 5' primer bind, sequence tagged site.
ACCESSION L30907
VERSION L30907.1 GI:624443
KEYWORDS SRS; PCR primer; SRS sequence; microsatellite DNA; microsatellite marker; sequence tagged site; tetranucleotide repeat.

SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 57)
Gerken,S.C., Matsunami,N., Plaetke,R., Albertsen,H., Ballard,L.,
Melis,R., Lawrence,E., Moore,M., Holik,P.R., Carlson,M., Zhao,X.,
Robertson,M., Bradley,P., Elsner,T., Tingey,A., Lalouel,J.-M. and
White,R.
Genetic and physical mapping of simple sequence repeat containing
sequence tagged sites from the human genome
Unpublished (1994)
Submitted by: Utah Center for Human Genome Research University of
Utah, Dept. of Human Genetics
2160 Eccles Institute of Human Genetics
Salt Lake City, UT 84112
e-mail: steecorona.med.utah.edu
Primer A: CGAGCTCCGTCACAGAAA
Primer B: CCATCCACTCTTCATG
End to Label: Primer A
PCR Profile:
Initial Denaturation: 94C 300sec
Cycles Denaturation Annealing Extension 5 94
C 10 sec. 56 C 10 sec. 72 C 20 sec. 30 94 C 10 sec.
56 C 10 sec. 72 C 20 sec. Mg++: 1.00 mM
Gel: Acrylamide 7%, Formamide 32%, Urea 34%
Alleles: 5.

FEATURES
source 1..57
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
32..50
primer_bind /evidence=experimental 7 t
BASE COUNT 21 a 16 c 13 g 7 t
ORIGIN

Query Match 0.1%; Score 36; DB 13; Length 57;
Best Local Similarity 100.0%; Pred. No. 3.1e-07;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8625 GAGCTCGCTGTGACCCAGCGTGGAGTGCATGG 8660
DB 39 GAGCTCGCTGTGACCCAGCGTGGAGTGCATGG 4

RESULT 6
AF087511 66 bp mRNA PRI 13-SEP-1999
LOCUS AF087511 66 bp mRNA PRI 13-SEP-1999
DEFINITION Homo sapiens clone ENAc+22 epithelial sodium channel subunit
(SCNN1A) mRNA, alternatively spliced, partial sequence.
ACCESSION AF087511

VERSION AF087511.1 GI:5870626
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 66)
Oh,Y. and Warnock,D.G.
An Alu cassette in the human epithelial sodium channel
Unpublished
2 (bases 1 to 66)
Oh,Y. and Warnock,D.G.
Direct Submission
Submitted (26-AUG-1998) Medicine, UAB, 1720 7th Ave. SO.,
Birmingham, AL 35294, USA
JOURNAL

FEATURES
source 1..66
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="12"
/clone="ENAc+22"
<1..>66
/gene="SCNN1A"
/note="alternatively spliced epithelial sodium channel
alpha subunit containing intronic Alu sequence"
<1..>66
/gene="SCNN1A"

BASE COUNT 12 a 20 c 24 g 10 t
ORIGIN

Query Match 0.1%; Score 36; DB 40; Length 66;
Best Local Similarity 100.0%; Pred. No. 3.1e-07;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28853 CCGCCTCGGCTCCCAAGTGTGGATTACAGGT 28888
DB 55 CCGCCTCGGCTCCCAAGTGTGGATTACAGGT 20

RESULT 7
HUMUT1543A 57 bp DNA SRS 17-JAN-1995
LOCUS HUMUT1543A 57 bp DNA SRS 17-JAN-1995
DEFINITION Human SRS UT1543, 5' primer bind, sequence tagged site.
ACCESSION L30907
VERSION L30907.1 GI:624443
KEYWORDS SRS; PCR primer; SRS sequence; microsatellite DNA; microsatellite marker; sequence tagged site; tetranucleotide repeat.

SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 57)
Gerken,S.C., Matsunami,N., Plaetke,R., Albertsen,H., Ballard,L.,
Melis,R., Lawrence,E., Moore,M., Holik,P.R., Carlson,M., Zhao,X.,
Robertson,M., Bradley,P., Elsner,T., Tingey,A., Lalouel,J.-M. and
White,R.
Genetic and physical mapping of simple sequence repeat containing
sequence tagged sites from the human genome
Unpublished (1994)
Submitted by: Utah Center for Human Genome Research University of
Utah, Dept. of Human Genetics
2160 Eccles Institute of Human Genetics
Salt Lake City, UT 84112
e-mail: steecorona.med.utah.edu
Primer A: CGAGCTCCGTCACAGAAA
Primer B: CCATCCACTCTTCATG
End to Label: Primer A
PCR Profile:
Initial Denaturation: 94C 300sec
Cycles Denaturation Annealing Extension 5 94
C 10 sec. 56 C 10 sec. 72 C 20 sec. 30 94 C 10 sec.
56 C 10 sec. 72 C 20 sec. Mg++: 1.00 mM

TITLE
JOURNAL
COMMENT

Gel: Acrylamide 7%, Formamide 32%, Urea 34%
Alleles: 5.
FEATURES
source location/Qualifiers
1. .57
/organism="Homo sapiens"
/db_xref="taxon:9606"
primer_bind 32. .50
/evidence=experimental
BASE COUNT 21 a 16 c 13 g 7 t
ORIGIN

Query Match 0.1%; Score 35; DB 13; Length 57;
Best Local Similarity 100.0%; Pred. No. 1.1e-06;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 26376 GCATCCAGCCTGGGTGACGAGCGAGACTCCGTC 26410
Db 9 GCATCCAGCCTGGGTGACGAGCGAGACTCCGTC 43

RESULT 8
LOCUS AR051522 85 bp DNA PAT 29-SEP-1999
DEFINITION Sequence 92 from patent US 5830670.
ACCESSION AR051522
VERSION AR051522.1 GI:5974886
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 85)
AUTHORS de la Monte, S. and Wands, J.R.
TITLE Neural thread protein gene expression and detection of Alzheimer's disease
JOURNAL Patent: US 5830670-A 92 03-NOV-1998;
FEATURES Location/Qualifiers
source 1. .85
/organism="unknown"
BASE COUNT 13 a 27 c 25 g 20 t
ORIGIN

Query Match 0.1%; Score 35; DB 5; Length 85;
Best Local Similarity 100.0%; Pred. No. 1.1e-06;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 21802 CGGGCGATCAGAGGTGAGATCGAGATCGACATC 21836
Db 48 CGGGCGATCAGAGGTGAGATCGAGATCGACATC 14

RESULT 9
LOCUS HDMALVANC 42 bp DNA PRI 08-OCT-1994
DEFINITION Homo sapiens 4000 year old remains from Nekht-anhk Alu repeat fragment 12:4.
ACCESSION L36838
VERSION L36838.1 GI:556195
KEYWORDS Alu repeat.
SOURCE Homo sapiens (individual isolate 4000 year old remains from Nekht-anhk) liver DNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (sites)
AUTHORS Pabo, S.
TITLE Ancient DNA: extraction, characterization, molecular cloning, and enzymatic amplification
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 86 (6), 1939-1943 (1989)
FEATURES
source location/Qualifiers
1. .42

/organism="Homo sapiens"
/isolate="4000 year old remains from Nekht-anhk"
/db_xref="taxon:9606"
/tissue_type="liver"
repeat_region <1. .742
/rpt_family="Alu"
BASE COUNT 8 a 10 c 17 g 7 t
ORIGIN

Query Match 0.1%; Score 34; DB 9; Length 42;
Best Local Similarity 100.0%; Pred. No. 3.5e-06;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 28846 TGATCCACCGCTGCGCTCCCAAGTGTGG 28879
Db 34 TGATCCACCGCTGCGCTCCCAAGTGTGG 1

RESULT 10
LOCUS AR051521 84 bp DNA PAT 29-SEP-1999
DEFINITION Sequence 91 from patent US 5830670.
ACCESSION AR051521
VERSION AR051521.1 GI:5974885
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 84)
AUTHORS de la Monte, S. and Wands, J.R.
TITLE Neural thread protein gene expression and detection of Alzheimer's disease
JOURNAL Patent: US 5830670-A 91 03-NOV-1998;
FEATURES Location/Qualifiers
source 1. .84
/organism="unknown"
BASE COUNT 15 a 29 c 21 g 19 t
ORIGIN

Query Match 0.1%; Score 34; DB 5; Length 84;
Best Local Similarity 100.0%; Pred. No. 3.6e-06;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2676 GCCCGCTAGCCTCCCAAGTGTGGATTACA 2709
Db 44 GCCCGCTAGCCTCCCAAGTGTGGATTACA 77

RESULT 11
LOCUS A25212 35 bp DNA PAT 11-APR-1995
DEFINITION Inter-Alu specific primer DNA (pdj33) from patent WO9213101.
ACCESSION A25212
VERSION A25212.1 GI:904592
KEYWORDS synthetic construct.
SOURCE synthetic construct.
ORGANISM artificial sequence.
REFERENCE 1 (bases 1 to 35)
AUTHORS
TITLE METHOD OF DETECTING DNA SEQUENCE VARIATION
JOURNAL Patent: WO 9213101-A 3 06-AUG-1992;
FEATURES Location/Qualifiers
source 1. .35
/organism="synthetic construct"
/db_xref="taxon:32630"
BASE COUNT 8 a 10 c 11 g 6 t
ORIGIN

Query Match 0.1%; Score 33; DB 5; Length 35;

Best Local Similarity 100.0%; Pred. No. 1.2e-05;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1886 TGCGTCACGCCCTGTAATCCAGCAGCACTTTGGAG 1918
Db 35 TGCGTCACGCCCTGTAATCCAGCAGCACTTTGGAG 3

RESULT 12

E09140/c
ID E09140 standard; DNA; UNC; 35 BP.

AC E09140;

XX E09140.1

DT 07-OCT-1997 (Rel. 52, Created)

DT 08-OCT-1997 (Rel. 52, Last updated, Version 1)

DE Synthetic DNA for Alu specific primer.

XX JP 1995115999-A/3.

OS unidentified

OC unclassified.

XX [1]

RP 1-35

RA Andoreasu H.A., Yan F.;

RT "DETECTING METHOD FOR DNA ARRANGEMENT VARIATION";

RL Patent number JP 1995115999-A/3, 09-MAY-1995.

RL INGENI BV.

XX OS None

CC OC Artificial sequences.

CC PN JP 1995115999-A/3

CC PD 09-MAY-1995

CC PF 22-MAY-1992 JP 1992130668

CC PI ANDOREASU HERARUDOUSU AITSUTERURINDEN, YAN FUEIKU

CC PC C1201/68, C12N15/00, G01N27/447, G01N27/447;

CC CC strandedness: Single;

CC CC topology: Linear;

CC CC Key Location/Qualifiers

CC FT source 1..35

CC FT misc-feature 1..35

CC FT misc-feature 1..35

CC FT misc-feature 1..35

CC FT misc-feature 1..35

CC FT misc-feature 1..35

CC FT misc-feature 1..35

CC FT misc-feature 1..35

CC FT misc-feature 1..35

CC FT misc-feature 1..35

CC FT misc-feature 1..35

CC FT misc-feature 1..35

CC FT misc-feature 1..35

CC FT misc-feature 1..35

CC FT misc-feature 1..35

CC FT misc-feature 1..35

CC FT misc-feature 1..35

CC FT misc-feature 1..35

CC FT misc-feature 1..35

CC FT misc-feature 1..35

CC FT misc-feature 1..35

CC FT misc-feature 1..35

CC FT misc-feature 1..35

CC FT misc-feature 1..35

CC FT misc-feature 1..35

CC FT misc-feature 1..35

CC FT misc-feature 1..35

CC FT misc-feature 1..35

CC FT misc-feature 1..35

CC FT misc-feature 1..35

CC FT misc-feature 1..35

KEYWORDS
SOURCE unidentified.

ORGANISM unidentified.

REFERENCE 1 (bases 1 to 40)

AUTHORS Resnick M.A., Larionov V.L., Koupina N.Y. and Perkins E.L.

TITLE TRANSFORMATION-ASSOCIATED RECOMBINATION CLONING

JOURNAL Patent: WO 9801573-A 15-JAN-1998;

US HEALTH (US)

FEATURES Location/Qualifiers

source 1..40

BASE COUNT 7 a 12 c 13 g 8 t

ORIGIN

Query Match 0.1%; Score 33; DB 5; Length 40;

Best Local Similarity 100.0%; Pred. No. 1.2e-05;

Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1886 TGCGTCACGCCCTGTAATCCAGCAGCACTTTGGAG 1918

Db 3 TGCGTCACGCCCTGTAATCCAGCAGCACTTTGGAG 35

RESULT 14
HUMALUANC
LOCUS HUMALUANC 65 bp DNA PRI 08-OCT-1994

DEFINITION Homo sapiens 4000 year old remains from Nekht-anhk Alu repeat 12:2.

ACCESSION L36836

VERSION L36836.1 GI:556194

KEYWORDS Homo sapiens (individual isolate 4000 year old remains from

source Nekht-anhk) liver DNA.

ORGANISM Homo sapiens

REFERENCE 1 (sites)

AUTHORS Paabo, S.

TITLE Ancient DNA: extraction, characterization, molecular cloning, and

JOURNAL enzymatic amplification

MEDLINE Proc. Natl. Acad. Sci. U.S.A. 86 (6), 1939-1943 (1989)

FEATURES 89184542

source Location/Qualifiers

1..65

/organism="Homo sapiens"

/isolate="4000 year old remains from Nekht-anhk"

/db_xref="taxon:9606"

/tissue-type="liver"

repeat-region <1..>65

BASE COUNT 17 a 17 c 21 g 10 t

ORIGIN

Query Match 0.1%; Score 33; DB 9; Length 65;

Best Local Similarity 100.0%; Pred. No. 1.2e-05;

Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 26168 TTGGAGAGCGCGAGCGGCGATCAGAGGTC 26200

Db 33 TTGGAGAGCGCGAGCGGCGGATCAGAGGTC 65

RESULT 15
HUMUT578A
LOCUS HUMUT578A 69 bp DNA SRS 19-JUL-1995

DEFINITION Human SRS UT578, 5' primer bind, sequence tagged site.

ACCESSION L39139 L18336

VERSION L39139.1 GI:642101

KEYWORDS SRS; PCR primer; SRS sequence; microsatellite DNA; microsatellite

marker; sequence tagged site; tetranucleotide repeat.

Mon Apr 3 08:24:28 2000

us-08-852-495c-1_copy_65000_95000.rge

Page 6

```
SOURCE      Homo sapiens DNA.
ORGANISM     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
              Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1 (bases 1 to 89)
AUTHORS      Gerken, S.C., Matsunami, N., Plaetke, R., Albertsen, H., Ballard, L.,
              Wells, R., Lawrence, E., Moore, M., Holik, P.R., Carlson, W., Zhao, X.,
              Robertson, M., Bradley, P., Eisner, T., Tingey, A., Lalouel, J.-M. and
              White, R.
TITLE        Genetic and physical mapping of simple sequence repeat containing
              sequence tagged sites from the human genome
JOURNAL      Unpublished (1994)
COMMENT      Submitted by: Utah Center for Human Genome Research University of
              Utah, Dept. of Human Genetics
              2160 Eccles Institute of Human Genetics
              Salt Lake City, UT 84112
              e-mail: sts@corona.med.utah.edu
              Primer A: AGTGCAGACCGCCTGGC
              Primer B: AGTGGCAGAAAATCGCATC
              End to Label: Primer A
PCR Profile:
  Initial Denaturation: 94C 300sec
  Cycles:
    Denaturation: 94C 30sec
    Annealing: 58C 10sec
    Extension: 72C 20sec
    Mg++: 1.50 mM
    gel: Acrylamide 7%, Formamide 32%, Urea 34%
  Alleles: 1
FEATURES     location/Qualifiers
             source          1..69
                               /organism="Homo sapiens"
                               /db_xref="taxon:9606"
                               /map="19"
             primer_bind     14..32
                               /note="STS UT578 5' end"
                               /evidence=experimental
             base_count       26 a 16 c 14 g 13 t
             origin           13 t

Query Match      0.1%; Score 33; DB 13; Length 69;
Best Local Similarity 100.0%; Pred. No. 1.2e-05;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Ox 15491 GTGGCCAGAGCTGCTGCACTCTGACCTCA 15523
Db 36 GTGGCCAGAGCTGCTGCACTCTGACCTCA 4

Search completed: March 30, 2000, 07:33:31
Job time: 101360 sec
```


KW co-polymer; sutures; drug delivery device; screws; ds.
 OS Synthetic.
 PN CA2112716-A.
 PD 01-JUL-1994.
 PF 31-DEC-1993; 112716.
 PR 31-DEC-1992; US-999520.
 PA (USSTU) US SURGICAL CORP.
 PI Gruskis EA;
 DR WPI: 94-272324/34.
 PT New biocompatible medical devices and articles - using a lactic acid-glycolic acid polyester in which monomer units are individually and specifically determined
 PS Example 1: Page 45; 67pp; English
 CC This sequence represents the 5' fragment of the synthetic genes which encode polyesters given in Q71801-02. This fragment was ligated into the Klenow blunted sequences given in Q71808-09 which represent the 3' and 5' ends of the plasmid pSPORT 1B digested with FokI. This places this synthetic fragment under the control of the E. coli trpA initiator. The E. coli trpA initiator sequence was used to express the sequences given in Q71801-02 in an E. coli cell free medium. The expressed sequences represent synthetic genes which were used in the synthesis of polyesters derived from the α -amino acid analogues of an amino acid except pro. Of the three stop codons, UAA, UAG and UGA, one of these stop codons is reversed for the stop signal for the polyester. The remaining two stop codons are available for encoding the monomer units of the polyester. Lactate and glycolate are encoded by UAA and UAG, respectively. Initiation of translation always starts at an AUG codon, therefore the first lactate of the chain may be encoded by an AUG codon. This is accomplished through the use of a specially modified Met-tRNA^{Met} modified to carry lactate instead of Met. Alternatively, a Met is incorporated in the first position of the polymer chain. The resulting polymer is treated with cyanogen bromide to remove the Met. Co-polymers produced in this manner may be used to produce sutures, staples, clips, drug delivery devices, pins and screws.
 Sequence 80 BP; 47 A; 6 C; 4 G; 23 T;
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492

CC	ribozymes and contains the necessary initiator sequences and the
CC	template for the polypeptide leader. The resulting plasmid pMAL-2c
CC	contains the complete polyester coding sequence ligated in frame with
CC	the leader sequence. The expressed sequences represent synthetic
CC	genes which were used in the synthesis of polyesters derived from the
CC	c amino acid analogues of an amino acid except Pro. Of the three stop
CC	codons, UAA, UAG and UGA, one of these stop codons is reversed for the
CC	stop signal for the polyester. The remaining two stop codons are
CC	available for encoding the monomer units of the polyester. Lactate
CC	and glycolate are encoded by UAA and UAG, respectively. Initiation of
CC	translation always starts at an AUG codon, therefore the first lactate
CC	of the chain may be encoded by an AUG codon. This is accomplished
CC	through the use of a specially modified Met-tRNA ^{fMet} modified to carry
CC	lactate instead of Met. Alternatively, a Met is incorporated in the
CC	first position of the polymer chain. The resulting polymer is treated
CC	with cyanogen bromide to remove the Met. Copolymers produced in this
CC	manner may be used to produce sutures, staples, clips, drug delivery
CC	devices, pins and screws.
CC	Sequence 81 bp; 6 C; 4 G; 24 T;
DG	
OY	2146 AATATAATATAATATAATATAATATAATATAATATAA 2186
DG	2 AATATAATATAATATAATATAATATAATATAATATAA 42
RESULT	
ID	071809
AC	G71809 standard; DNA; 103 BP.
DT	30-MAR-1995 (first entry)
DE	psport-ic synthetic polyester gene 5' fragment and trpa initiator.
KM	Synthetic gene; synthesis; polyester; a-amino acid analogue; clips;
KW	stop codon; stop signal; monomer unit; lactate; glycolate; staples;
KW	translation initiation; modified; Met-tRNA ^{fMet} ; cyanogen bromide; pins;
OS	co-polymer; sutures; drug delivery device; screws; ds.
FT	Key
FT	misc_signal
FT	1..18
FT	/tag= a
FT	/note= "trpa initiator"
FT	19..21
FT	/tag= b
FT	/note= "Start codon"
FT	22..93
FT	misc_feature
FT	22..93
FT	/tag= c
FT	/note= "Polyester gene 5' fragment"
PN	CNA112716-A.
PF	01-JUL-1994.
PF	31-DEC-1993; 112716.
PR	31-DEC-1992; US-999520.
PA	(USSV) US SURGICAL CORP.
DR	Gruskin EA.
PI	WPI; 94-27334/34.
PT	New biocompatible medical devices and articles - using a lactic
PT	acid glycolic acid polyester in which monomer units are
PT	individually and specifically determined
PS	Example 1; Page 45; 67pp; English.
CC	This sequence represents a fragment of the plasmid psport ic which
CC	contains the 5' fragment of the synthetic genes which encode polyesters,
CC	given in G71801-02, under the control of the E. coli trpa transcription
CC	initiator and followed by PstI and HindIII restriction sites. A further
CC	fragment of the synthetic genes were ligated into psport ic to give
CC	plasmid psport id (see also G71801-11). With the final fragment being
CC	inserted in a third round of ligation to give psport ie (see also G71812-
CC	13) The E. coli trpa initiator sequence was used to express the full
CC	length sequences given in G71801-02 in an E coli cell free medium. The
CC	expressed sequences represent synthetic genes which were used in the
CC	synthesis of polyesters derived from the a-amino acid analogues of an

CC amino acid except Pro. Of the three stop codons, UAA, UAG and UGA, one
CC of these stop codons is reversed for the stop signal for the polyester.
CC The remaining two stop codons are available for encoding the monomer.
CC units of the polyester. Lactate and glycolate are encoded by UAA and
CC UAG, respectively. Initiation of translation always starts at an AUG
CC codon, therefore the first lactate of the chain may be encoded by an
CC AUG codon. This is accomplished through the use of a specially modified
CC Met-tRNA modified to carry lactate instead of Met. Alternatively, a
CC Met is incorporated in the first position of the polymer chain. The
CC resulting polymer is treated with cyanogen bromide to remove the Met.
CC Co-polymers produced in this manner may be used to produce sutures,
CC staples, clips, drug delivery devices, pins and screws.
SQ Sequence 103 BP; 58 A; 7 C; 9 G; 29 T;

Query Match 0.1%; Score 41; DB 1; Length 103;
Best Local Similarity 100.0%; Pred. No. 0.0034;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2146 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2186
DB 23 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 63

RESULT 5
O27391/c
ID O27391 standard; DNA; 35 BP.
AC Q27391;
DE 27-JAN-1993 (first entry)
DE Inter-Alu specific primer PDJ33.
DE Polymerase chain reaction; PCR; repetitive element; ss.
OS Synthetic.
PN WD9213101-A.
PD 06-AUG-1992.
PF 24-JAN-1992; NL0018.
PR 25-JAN-1991; NL-000132.
PA (INGE-) INGENY BV.
PI Uiterlinden AG, Vijg J;
DR MPI: 92-286683/34.
PT Detection of genetic variation by 2-D electrophoresis of
PT fragments - and hybridisation with labelled probes, carried out
PT on fragments consisting of inter-repeat sequences generated by
PT PCR.
PS Claim 6; Page 6; 31pp; English.
CC Primer PDJ33 is one of several primers which are preferred for use
CC in amplifying inter-Alu regions of DNA. The amplified fragments are
CC then subjected to 2-D electrophoresis on the basis of length and
CC differences in base sequence. The resulting separation pattern is
CC transferred to a filter for screening with a probe. The method can
CC be used to detect genetic variation.
CC See Q27389-Q27404 and Q33141-Q33144.
SQ Sequence 35 BP; 8 A; 10 C; 11 G; 6 T;

Query Match 0.1%; Score 33; DB 1; Length 35;
Best Local Similarity 100.0%; Pred. No. 0.8;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1886 TGGCTCAGCGCTGTATCCAGCATTGGGAG 1918
DB 35 TGGCTCAGCGCTGTATCCAGCATTGGGAG 3

RESULT 6
V19044
ID V19044 standard; DNA; 40 BP.
AC V19044;
DE 28-JUL-1998 (first entry)
DE Alu PCR primer 1.
DE PCR; primer; amplification; Alu repeat sequence; vector;
KW circular yeast artificial chromosome; YAC; ss.
OS Synthetic.
OS Saccharomyces sp.

PN WO9801573-A1.
PD 15-JAN-1998.
PF 09-JUL-1996; U11478.
PR 09-JUL-1996; WO-U11478.
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
PI Kourina NY, Lariouev VL, Perkins EL, Resnick MA;
DR MPI: 98-110234/10.
PT Preparation of yeast artificial chromosomes - by in vivo
PT recombination using vector comprising yeast centromere, marker,
PT yeast telomere and nucleic acid for recombination
PS Example 1; Page 45; 117pp; English.
CC This is the nucleotide sequence for the PCR primer used in the
CC amplification of the Alu repeat sequence, which is used to
CC demonstrate the processes described in the invention. It involves
CC the creation and use of circular yeast artificial chromosome (YAC)
CC to selectively clone specific nucleic acids from a background of
CC mixed nucleic acids by introducing the vector(s) into E. coli cells.
CC They can be used to rapidly isolate human DNA where only a part of the
CC sequence of DNA is known. Using the methods large fragments of DNA can
CC be easily cloned and analysed.
SQ Sequence 40 BP; 7 A; 12 C; 13 G; 8 T;

Query Match 0.1%; Score 33; DB 1; Length 40;
Best Local Similarity 100.0%; Pred. No. 0.77;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1886 TGGCTCAGCGCTGTATCCAGCATTGGGAG 1918
DB 3 TGGCTCAGCGCTGTATCCAGCATTGGGAG 35

RESULT 7
Q30397/c
ID Q30397 standard; DNA; 36 BP.
AC Q30397;
DE 07-DEC-1992 (first entry)
DE Oligomer IAP322 for forming triplex with HUMINT02 target duplex.
DE Human leukocyte adhesion protein; p150,95 alpha subunit gene; herpes
KW Simplex; AIDS; modified; HIV; RSV; HPV; malignancy; hepatitis;
KW inflammation; ss.
OS Synthetic.
FH Key
FT modified_base 10 Location/Qualifiers
FT /*tag= a
FT /mod_base= OTHER
FT /note= "OTHER- N6 methyl-8-oxo 2' deoxyadenine"
FT modified_base 13
FT /*tag= b
FT /mod_base= OTHER
FT /note= "OTHER- N6 methyl-8-oxo 2' deoxyadenine"
FT modified_base 16
FT /*tag= c
FT /mod_base= OTHER
FT /note= "OTHER- N6 methyl-8-oxo 2' deoxyadenine"
FT modified_base 19
FT /*tag= d
FT /mod_base= OTHER
FT /note= "OTHER- N6 methyl-8-oxo 2' deoxyadenine"
FT modified_base 22
FT /*tag= e
FT /mod_base= OTHER
FT /note= "OTHER- N6 methyl-8-oxo 2' deoxyadenine"
FT modified_base 23
FT /*tag= f
FT /mod_base= OTHER
FT /note= "OTHER- N6 methyl-8-oxo 2' deoxyadenine"
FT modified_base 28
FT /*tag= g
FT /mod_base= OTHER
FT /note= "OTHER- N6 methyl-8-oxo 2' deoxyadenine"
FT modified_base 31
FT /*tag= h

```
FT /mod_base- OTHER
FM /note- OTHER- N6 methyl-8-oxo 2' deoxyadenine"
FT 344
FT /tag- 1
FT /mod_base- OTHER
FT /note- "OTHER- N6 methyl-8-oxo 2' deoxyadenine"
PN W09205705-A.
PD 11-JUN-1992.
PF 25-NOV-1991.
PR 23-NOV-1990; US-617907.
PR 18-JAN-1991; US-643382.
PR 08-APR-1991; US-683420.
PR 17-APR-1991; US-686547.
PR 17-APR-1991; US-686546.
PR 27-SEP-1991; US-766733.
PR (GILE-) GILEAD SCI INC.
PA Firehler B, Krawczyk S, Matteucci MD, Milligan J;
PI WPI: 92-217083/26.
PT New oligomers contg. modified bases - which form a triplex with
PT G-C doublet in a DNA duplex, for treating and diagnosing HIV,
PT hepatitis, herpes, malignant tumours and inflammation
PT Claim 12, Page 70; 77pp; English.
PS The synthetic oligomer is capable of forming a triplex at
CC physiological pH with a putine rich target sequence by coupling
CC into the major groove of the duplex. The specific target sequence
CC of this oligomer is the human leukocyte adhesion protein P150, 95
CC alpha subunit gene (HUMINTO2) beginning at nucleotide 2370 contg. a
CC putine rich sequence contd. on one strand of the duplex. The oligomer,
CC and others like it are useful in diagnosis and therapy of diseases
CC characterised by specific DNA duplex targets, e.g. HPV; HER; HIV;
CC hepatitis B; herpes; malignant tumours and inflammation. The triplex
CC helices form under mild conditions thus assays may be carried out
CC without subjecting the test specimen to harsh conditions. The
CC oligomer contains an inverted polarity region formed from an o-xyloso
CC dimer synthon. The linking sp. is o-xyloso (nucleosides have the 3'
CC positions of xylose sugars linked via the o-xyloso ring). Two
CC nucleotides are coupled through a xylose residue to form the dimer
CC synthon. This additional modifications may render the oligomer stable
CC to nuclease activity. The oligomer is able to inhibit gene expression,
CC as verified by in vitro systems.
CC See also Q25452-25501 and Q30226-448.
CC Sequence 36 BP; 9 A; 0 C; 0 G; 27 T;
SQ
```

```
Query Match 0.1%; Score 30; DB 1; Length 36;
Best Local Similarity 100.0%; Pred. No. 5.5;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 2158 AATTAATTAATTAATTAATTAATTAATTA 2187
DB 36 AATTAATTAATTAATTAATTAATTAATTA 7
```

```
RESULT 8
ID V19044/c
AC V19044;
DR 28-JUL-1998 (first entry)
DE Alu PCR primer 1.
KW PCR; primer; amplification; Alu repeat sequence; vector;
KV circular yeast artificial chromosome; YAC; ss.
OS Synthetic.
OS Saccharomyces sp.
PN W09801573-A1.
PD 15-JAN-1998.
PR 09-JUL-1996; U11478.
PR 09-JUL-1996; NO-U11478.
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
PA Koudrina NT, Laktionov VL, Perkins EL, Resnick MA;
DR WPI: 98-110234/10.
PT Preparation of yeast artificial chromosomes - by in vivo
PT recombination using vector comprising yeast centromere, marker,
```

```
PT Yeast telomere and nucleic acid for recombination
PS Example 1; Page 45; 117pp; English.
CC This is the nucleotide sequence for the PCR primer used in the
CC amplification of the Alu repeat sequence, which is used to
CC demonstrate the processes described in the invention. It involves
CC the creation and use of circular yeast artificial chromosome (YAC)
CC to selectively clone specific nucleic acids from a background of
CC mixed nucleic acids by introducing the vector(s) into E. coli cells.
CC They can be used to rapidly isolate human DNA where only a part of the
CC sequence of DNA is known. Using the methods large fragments of DNA can
CC be easily cloned and analysed.
SQ Sequence 40 BP; 7 A; 12 C; 13 G; 8 T;
```

```
Query Match 0.1%; Score 30; DB 1; Length 40;
Best Local Similarity 100.0%; Pred. No. 5.4;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 8859 AAGTGTGGGATTACAGCGTGAGCCACC 8888
DB 30 AAGTGTGGGATTACAGCGTGAGCCACC 1
```

```
RESULT 9
ID T26410/c
AC T26410;
DR 06-DEC-1996 (first entry)
DE Human gene signature HUMGS08651.
KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.
OS Homo sapiens.
PN W09514772-A1.
PD 01-JUN-1995.
PF 11-NOV-1994; J01916.
PR 12-NOV-1993; JP-355504.
PA (MATS/) MATSUBARA K.
PA (OKUBA/) OKUBO K.
PI Matsubara K, Okubo K;
DR WPI: 95-206931/27.
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function. By preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues
PS Claim 1; Page 2078; 2245pp; Japanese.
CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in T19001-T26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.
SQ Sequence 91 BP; 38 A; 22 C; 11 G; 17 T;
```

```
Query Match 0.1%; Score 30; DB 1; Length 91;
Best Local Similarity 100.0%; Pred. No. 4.4;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 2460 AGCTGTGCTGTGTTGCCAGGCTGGAGTGC 2489
DB 42 AGCTGTGCTGTGTTGCCAGGCTGGAGTGC 13
```

```
RESULT 10
ID 027391 standard; DNA; 35 BP.
AC 027391;
DT 27-JAN-1993 (first entry)
DE Inter-Alu specific primer PDJ33.
KM Polymerase chain reaction; PCR; repetitive element; ss.
OS Synthetic.
PN W09213101-A.
PD 06-AUG-1992.
PF 24-JAN-1992; NL0018.
PR 25-JAN-1991; NL-000132.
PA (INGE-) INGENY BV.
PI Uiterlinden AG, Vljg J;
DR WPI; 92-284683/34.
PT Detection of genetic variation by 2-D electrophoresis of
PT fragments - and hybridisation with labelled probes, carried out
PT on fragments consisting of inter-repeat sequences generated by
PT PCR.
PS Claim 6; Page 6; 31pp; English.
CC Primer PDJ33 is one of several primers which are preferred for use
CC in amplifying inter-Alu regions of DNA. The amplified fragments are
CC then subjected to 2-D electrophoresis on the basis of length and
CC differences in base sequence. The resulting separation pattern is
CC transferred to a filter for screening with a probe. The method can
CC be used to detect genetic variation.
CC See 027389-027404 and 033141-033144.
SQ Sequence 35 BP; 8 A; 10 C; 11 G; 6 T;

Query Match 0.1%; Score 28; DB 1; Length 35;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8859 AAAGTCTGGGATTACAGCGTGAGCCA 8886
DB 8 AAAGTCTGGGATTACAGCGTGAGCCA 35
|||||
V83640/c
ID V83640 standard; DNA; 46 BP.
AC V83640;
DT 01-MAR-1999 (first entry)
DE Oligonucleotide 1 used in the construction of assay plasmids.
KM Repetitive sequence; carcinogenic; human dietary component;
KM DNA instability; cancer; diet; primer; ss.
OS Synthetic.
PN W09845476-A1.
PD 15-OCT-1998.
PF 08-APR-1998; G00869.
PR 08-APR-1997; GB-007141.
PA (FOOD-) FOOD RES INST.
PI Schweizer M;
DR WPI; 99-024011/02.
PT Assay for testing the carcinogenic properties of a test substance
PT by introduction of a reporter gene expression vector containing a
PT repetitive DNA sequence that is unstable in cancer cells
PS Disclosure; Page 15; 103pp; English.
CC The present sequence represents an oligonucleotide used in the
CC construction of assay plasmids, which are used in the course of the
CC invention. The specification describes an assay for testing the
CC carcinogenic properties of a test substance. The assay comprises
CC introducing into cells a reporter gene expression vector comprising a
CC repetitive DNA sequence which exhibits instability in cancer cells,
CC whereby instability of the repetitive DNA sequence affects expression of
CC the reporter gene, exposing the resulting cells to the test substance
CC and determining whether the test substance is carcinogenic or
CC anti-carcinogenic by comparing the frequency of reporter gene expression
CC in the resulting cells with the frequency of reporter gene expression in
CC cells which have not been exposed to the test substance. The assay can
CC be used to identify human dietary components that protect against DNA
CC instability, and therefore some types of cancer, and can be used to
```

```
CC contribute to the scientific basis for a healthy diet.
SQ Sequence 46 BP; 18 A; 20 C; 6 G; 2 T;

Query Match 0.1%; Score 28; DB 1; Length 46;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18617 GCTGTGTGTGTGTGTGTGTGTGTG 18644
DB 40 GCTGTGTGTGTGTGTGTGTGTGTG 13
|||||
V83641
ID V83641 standard; DNA; 46 BP.
AC V83641;
DT 01-MAR-1999 (first entry)
DE Oligonucleotide 2 used in the construction of assay plasmids.
KM Repetitive sequence; carcinogenic; human dietary component;
KM DNA instability; cancer; diet; primer; ss.
OS Synthetic.
PN W09845476-A1.
PD 15-OCT-1998.
PF 08-APR-1998; G00869.
PR 08-APR-1997; GB-007141.
PA (FOOD-) FOOD RES INST.
PI Schweizer M;
DR WPI; 99-024011/02.
PT Assay for testing the carcinogenic properties of a test substance
PT by introduction of a reporter gene expression vector containing a
PT repetitive DNA sequence that is unstable in cancer cells
PS Disclosure; Page 16; 103pp; English.
CC The present sequence represents an oligonucleotide used in the
CC construction of assay plasmids, which are used in the course of the
CC invention. The specification describes an assay for testing the
CC carcinogenic properties of a test substance. The assay comprises
CC introducing into cells a reporter gene expression vector comprising a
CC repetitive DNA sequence which exhibits instability in cancer cells,
CC whereby instability of the repetitive DNA sequence affects expression of
CC the reporter gene, exposing the resulting cells to the test substance
CC and determining whether the test substance is carcinogenic or
CC anti-carcinogenic by comparing the frequency of reporter gene expression
CC in the resulting cells with the frequency of reporter gene expression in
CC cells which have not been exposed to the test substance. The assay can
CC be used to identify human dietary components that protect against DNA
CC instability, and therefore some types of cancer, and can be used to
CC contribute to the scientific basis for a healthy diet.
SQ Sequence 46 BP; 2 A; 6 C; 20 G; 18 T;

Query Match 0.1%; Score 28; DB 1; Length 46;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18617 GCTGTGTGTGTGTGTGTGTGTGTG 18644
DB 11 GCTGTGTGTGTGTGTGTGTGTGTG 38
|||||
V83635
ID V83635 standard; DNA; 47 BP.
AC V83635;
DT 01-MAR-1999 (first entry)
DE Oligonucleotide 1 used in the construction of assay plasmids.
KM Repetitive oligonucleotide; carcinogenic; human dietary component;
KM DNA instability; cancer; diet; ss.
OS Synthetic.
PN W09845476-A1.
PD 15-OCT-1998.
PF 08-APR-1998; G00869.
PR 08-APR-1997; GB-007141.
```

PA (FOOD-) FOOD RES INSTR.
P1 Schweizer M;
DR MPI; 99-024011/02.
PT Assay for testing the carcinogenic properties of a test substance
PT by introduction of a reporter gene expression vector containing a
repetitive DNA sequence that is unstable in cancer cells
PS Disclosure: Page 14; 103pp; English.
CC The present sequence represents a repetitive oligonucleotide used in the
CC construction of assay plasmids, which are used in the course of the
CC invention. The specification describes an assay for testing the
CC carcinogenic properties of a test substance. The assay comprises
CC introducing into cells a reporter gene expression vector carrying a
repetitive DNA sequence of the repetitive DNA sequence affects expression
CC whereby instability of the repetitive DNA sequence affects expression
CC of the reporter gene, exposing the test substance to the test substance
CC and determining whether the test substance is carcinogenic or
CC anti-carcinogenic by comparing the frequency of reporter gene expression
CC in the resulting cells with the frequency of reporter gene expression in
CC cells which have not been exposed to the test substance. The assay can
CC be used to identify human dietary components that protect against DNA
CC instability, and therefore some types of cancer, and can be used to
CC contribute to the scientific basis for a healthy diet.
SQ Sequence 47 Bp; 2 A; 6 C; 21 G; 18 T;

```
QY 18617 GCTGTGTCGTGTCGTGTCGTGTCGTG 18644
      |||
Db    11 GCTGTGTGTGTGTGTGTGTGTGTGTGTG 38

Query Match          0.1%; Score 28; DB 1; Length 47;
Best Local Similarity 100.0%; Pred. No. 19;
Matches   28; Conservative 0; Mismatches 0; Indels 0; Gaps
```

RESULT 14	
095177	
ID	095177 standard; DNA; 57 BP.
AC	095177:
DN	08-FEB-1996 (first entry)
DE	Simple tandem repeat (STR) corresponding to the 2nd part of wgl2.
KW	Simple tandem repeat; STR; wgl2; treatment; genetic; diagnosis;
KW	characterisation; mapping; linkage studies; analysts; alleles;
KW	second part; ss.
OS	Synthetic.
PN	W09517522-A2.
PD	29-JUN-1995.
PR	21-DEC-1994; G02789.
PR	21-DEC-1993; GB-026052.
PA	(UYLE-) UNIV LEICESTER.
P1	Armour J, Jeffreys AJ;
PI	WPI; 95-240682/31.
PT	Identifying simple tandem repeat loci in DNA - by screening DNA
PT	library to enrich for fragments contg. the repeats before cloning
PT	and rescreening, also simple tandem repeats for treatment or
PT	diagnosis
PS	Claim 26: Page 17: 51BP. English.
CC	095177 is a simple tandem repeat (SMR) corresponding to the 2nd part
CC	of wgl2. The SMR can be used for treatment and diagnosis in human
CC	and veterinary medicine, partic. for genetic characterisation,
CC	mapping, linkage studies and analysis/diagnosis of acquired disease
CC	alleles.
SO	Sequence 57 BP; 7 A; 11 C; 18 G; 21 T;

```

Query Match:          0.1%; Score 28; DB 1; Length 57;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0
QY 2467 CTCTGTGCCCAAGCTGAGTGCAGTGG 2484
      |||||
Db 23 CTCTGTGCCCAAGCTGAGTGCAGTGG 50

```

RESULT 15
Q33639 033639 standard; DNA; 64 BP.
AC 033639;
DT 02-FEB-1993 (first entry)
DE Microsatellite sequence from clone MTG13a.
KW PCR: selection; primers: OPIPR1M; breeding; cattle; parentage;
KW genetic mapping; traits; amplification; ss.
OS Bos taurus.
PN W09213102-A.
PD 05-AUG-1992.
FE 15-JAN-1992; U00340.
FR 15-JAN-1991; US-642342.
FA (GENM-) GENMARK.
FI Georges M, Massey JM;
DI WJ; 92-284664/34.
DR Polymorphic bovine DNA markers - used in genetic identification,
PT gene mapping, and selective breeding
PS Table 7, page 185; 517p; English.
CC The sequence is that of a bovine microsatellite sequence obtd. by
CC screening a library of bovine MbOI DNA fragments of between
CC 250 and 500 bp with an (AC)15 and a (TC)15 oligonucleotide probe.
CC One out of 50 clones cross-hybridised. Assuming independent
CC distribution of microsatellites and MbOI sites, the frequency of
CC (TG)n >9 microsatellites in the bovine genome is estimated at >100,
CC 000. The sequence information for ca. 230 such bovine microsatel-
CC les is summarised in the specification and indexed herein (see below).
CC The sequences upstream and downstream of the microsatellite sequence
CC were used to generate the required PCR primers for in vitro
CC amplification of the corresp. microsatellite (using the program
CC OPIPR1M). The microsatellites may be used to identify individuals,
CC for parentage testing, and in the genetic mapping of economic trait
CC loci, or genes involved, the determinism of economically important
CC traits esp. in cattle, to allow selective breeding.
CC See also Q33501-34437.
SQ Sequence 64 BP; 1 A; 2 C; 31 G; 30 T;

```
Query Match      0.1% Score 28; DB 1; Length 64;
Best Local Similarity 100.0%; Prod. No. 17;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0
```

OY 18619 TGTGTTGTGTGTGTGTGTGTGTGTCA 18646
|||||
Db 15 TGTGTGTGTGTGTGTGTGTGTGTCTCA 42

Search completed: March 30, 2000, 10:53:10
Job time: 111788 sec

Mon Apr 3 08:24:29 2000

us-08-852-495c-1_copy_65000_95000.png

Page 7



•

5

Mon Apr 3 08:24:30 2000

us-08-852-495c-1_copy_65000_95000.rml

Page 2

```
RESULT 2
US-08-340-426D-69
Sequence 69, Application US/08340426D
Patent No. 5948634
GENERAL INFORMATION:
APPLICANT: de la Monte, Suzanne
TITLE OF INVENTION: Neutral Thread Protein Gene Expression and Detection
NUMBER OF SEQUENCES: 121
CORRESPONDENCE ADDRESS:
ADDRESS: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/340,426D
FILING DATE: 14-NOV-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609.3840002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 69:
SEQUENCE CHARACTERISTICS:
LENGTH: 76 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
US-08-340-426D-69

Query Match 0.2% Score 58; DB 4; Length 76;
Best Local Similarity 100.0%; Pred. No. 2.5e-10;
Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 27685 GCCCGCGTAATTTTGTATTTAGTAGAGACAGGTTTCACCGTGGCCAGAGATG 27742
DB 6 GCCCGCGTAATTTTGTATTTAGTAGAGACAGGTTTCACCGTGGCCAGAGATG 63

RESULT 3
US-08-450-673C-69
Sequence 69, Application US/08450673C
Patent No. 5948888
GENERAL INFORMATION:
APPLICANT: de la Monte, Suzanne
TITLE OF INVENTION: Neutral Thread Protein Gene Expression and Detection
NUMBER OF SEQUENCES: 121
CORRESPONDENCE ADDRESS:
ADDRESS: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
```

```
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/450,673C
FILING DATE: 30-MAY-1995
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609.3840004
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 69:
SEQUENCE CHARACTERISTICS:
LENGTH: 76 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
US-08-450-673C-69

Query Match 0.2% Score 58; DB 4; Length 76;
Best Local Similarity 100.0%; Pred. No. 2.5e-10;
Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 27685 GCCCGCGTAATTTTGTATTTAGTAGAGACAGGTTTCACCGTGGCCAGAGATG 27742
DB 6 GCCCGCGTAATTTTGTATTTAGTAGAGACAGGTTTCACCGTGGCCAGAGATG 63

RESULT 4
PCT-US95-17111A-69
Sequence 69, Application PC/TUS9517111A
GENERAL INFORMATION:
APPLICANT: de la Monte, Suzanne
TITLE OF INVENTION: Neutral Thread Protein Gene Expression and Detection
NUMBER OF SEQUENCES: 121
CORRESPONDENCE ADDRESS:
ADDRESS: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/17111A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/340,426
FILING DATE: 14-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609.3840002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 69:
SEQUENCE CHARACTERISTICS:
LENGTH: 76 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
PCT-US95-17111A-69
```

Query Match 0.2%; Score 58; DB 6; Length 76;
Best Local Similarity 100.0%; Pred. No. 2.5e-10;
Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 27685 GCCCCGCTAATTTTGTATTTTGTAGAGACAGAGGTTTCACCGTGTGGCCAGAG 27742
|||||
Db 6 GCCCCGCTAATTTTGTATTTTGTAGAGACAGAGGTTTCACCGTGTGGCCAGAG 63

RESULT 5

US-08-454-557C-57
Sequence 57, Application US/08454557C
Patent No. 5830670
GENERAL INFORMATION:
APPLICANT: de la Monte, Suzanne
APPLICANT: Wands, Jack R.
TITLE OF INVENTION: Neutral Thread Protein Gene Expression and Detection
TITLE OF INVENTION: of Alzheimer's Disease
NUMBER OF SEQUENCES: 121
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/454, 557C
FILING DATE: 30-MAY-1995
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609.3840003
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ. ID NO: 57:
SEQUENCE CHARACTERISTICS:
LENGTH: 60 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
US-08-454-557C-57

Query Match 0.2%; Score 55; DB 3; Length 60;
Best Local Similarity 100.0%; Pred. No. 2.4e-09;
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 27685 GCCCCGCTAATTTTGTATTTTGTAGAGACAGAGGTTTCACCGTGTGGCCAGG 27739
|||||
Db 6 GCCCCGCTAATTTTGTATTTTGTAGAGACAGAGGTTTCACCGTGTGGCCAGG 60

RESULT 6

US-08-340-426D-57
Sequence 57, Application US/08340426D
Patent No. 5948634
GENERAL INFORMATION:
APPLICANT: de la Monte, Suzanne
APPLICANT: Wands, Jack R.
TITLE OF INVENTION: Neutral Thread Protein Gene Expression and Detection
TITLE OF INVENTION: of Alzheimer's Disease
NUMBER OF SEQUENCES: 121
CORRESPONDENCE ADDRESS:

ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/340, 426D
FILING DATE: 14-NOV-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609.3840002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ. ID NO: 57:
SEQUENCE CHARACTERISTICS:
LENGTH: 60 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
US-08-340-426D-57

Query Match 0.2%; Score 55; DB 4; Length 60;
Best Local Similarity 100.0%; Pred. No. 2.4e-09;
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 27685 GCCCCGCTAATTTTGTATTTTGTAGAGACAGAGGTTTCACCGTGTGGCCAGG 27739
|||||
Db 6 GCCCCGCTAATTTTGTATTTTGTAGAGACAGAGGTTTCACCGTGTGGCCAGG 60

RESULT 7

US-08-450-673C-57
Sequence 57, Application US/08450673C
Patent No. 5948888
GENERAL INFORMATION:
APPLICANT: de la Monte, Suzanne
APPLICANT: Wands, Jack R.
TITLE OF INVENTION: Neutral Thread Protein Gene Expression and Detection
TITLE OF INVENTION: of Alzheimer's Disease
NUMBER OF SEQUENCES: 121
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/450, 673C
FILING DATE: 30-MAY-1995
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609.3840004
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540

Mon Apr 3 08:24:30 2000

us-08-852-495c-1_copy_65000_95000.rni

Page 4

INFORMATION FOR SEQ ID NO: 57:
SEQUENCE CHARACTERISTICS:
LENGTH: 60 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
US-08-450-673c-57

Query Match 0.2%; Score 55; DB 4; Length 60;
Best Local Similarity 100.0%; Pred. No. 2.4e-09;
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 27685 GCGCCGCTATTTTGTATTTTGTAGTAGACAGAGGTTTCACCGTTGGCCAGG 27739
DB 6 GCGCCGCTATTTTGTATTTTGTAGTAGACAGAGGTTTCACCGTTGGCCAGG 60

RESULT 8
PCT-US95-17111A-57
Sequence 57, Application PC/TUS9517111A
GENERAL INFORMATION:
APPLICANT: de la Monte, Suzanne
TITLE OF INVENTION: Neural Thread Protein Gene Expression and
TITLE OF INVENTION: Detection of Alzheimer's Disease
NUMBER OF SEQUENCES: 121
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/17111A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/340,426
FILING DATE: 14-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609,3840002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2540
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 57:
SEQUENCE CHARACTERISTICS:
LENGTH: 60 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
PCT-US95-17111A-57

Query Match 0.2%; Score 55; DB 6; Length 60;
Best Local Similarity 100.0%; Pred. No. 2.4e-09;
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 27685 GCGCCGCTATTTTGTATTTTGTAGTAGACAGAGGTTTCACCGTTGGCCAGG 27739
DB 6 GCGCCGCTATTTTGTATTTTGTAGTAGACAGAGGTTTCACCGTTGGCCAGG 60

RESULT 9
US-08-454-557c-92/c

Sequence 92, Application US/08454557C
Patent No. 5830670
GENERAL INFORMATION:
APPLICANT: de la Monte, Suzanne
TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection
TITLE OF INVENTION: of Alzheimer's Disease
NUMBER OF SEQUENCES: 121
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/454,557C
FILING DATE: 30-MAY-1995
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609,3840003
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2540
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 92:
SEQUENCE CHARACTERISTICS:
LENGTH: 85 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
US-08-454-557c-92

Query Match 0.1%; Score 35; DB 3; Length 85;
Best Local Similarity 100.0%; Pred. No. 0.0071;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21802 CGGCGGATCAGAGGTGAGGATCGAGACATC 21836
DB 48 CGGCGGATCAGAGGTGAGGATCGAGACATC 14

RESULT 10
US-08-340-426D-92/c
Sequence 92, Application US/08340426D
Patent No. 5948631
GENERAL INFORMATION:
APPLICANT: de la Monte, Suzanne
TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection
TITLE OF INVENTION: of Alzheimer's Disease
NUMBER OF SEQUENCES: 121
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/340,426D

FILING DATE: 14-NOV-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609.3840002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 92:
SEQUENCE CHARACTERISTICS:
LENGTH: 85 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
US-08-340-426D-92

Query Match 0.1%; Score 35; DB 4; Length 85;
Best Local Similarity 100.0%; Pred. No. 0.0071;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21802 CGGGCGATCAGAGTCAAGATCGAGACATC 21836
|||||
Db 48 CGGGCGATCAGAGTCAAGATCGAGACATC 14

RESULT 11
US-08-450-673C-92/c
Sequence 92, Application US/08450673C
Patent No. 5948888
GENERAL INFORMATION:
APPLICANT: de la Monte, Suzanne
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609.3840004
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 92:
SEQUENCE CHARACTERISTICS:
LENGTH: 85 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
US-08-450-673C-92

Query Match 0.1%; Score 35; DB 4; Length 85;
Best Local Similarity 100.0%; Pred. No. 0.0071;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21802 CGGGCGATCAGAGTCAAGATCGAGACATC 21836
|||||
Db 48 CGGGCGATCAGAGTCAAGATCGAGACATC 14

RESULT 12
PCT-US95-17111A-92/c
Sequence 92, Application PC/TUS9517111A
GENERAL INFORMATION:
APPLICANT: de la Monte, Suzanne
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609.3840002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 92:
SEQUENCE CHARACTERISTICS:
LENGTH: 85 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
PCT-US95-17111A-92

Query Match 0.1%; Score 35; DB 6; Length 85;
Best Local Similarity 100.0%; Pred. No. 0.0071;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21802 CGGGCGATCAGAGTCAAGATCGAGACATC 21836
|||||
Db 48 CGGGCGATCAGAGTCAAGATCGAGACATC 14

RESULT 13
US-08-454-557C-91
Sequence 91, Application US/08454557C
Patent No. 5830670
GENERAL INFORMATION:
APPLICANT: de la Monte, Suzanne
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609.3840002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 92:
SEQUENCE CHARACTERISTICS:
LENGTH: 85 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
US-08-454-557C-91

Query Match 0.1%; Score 35; DB 6; Length 85;
Best Local Similarity 100.0%; Pred. No. 0.0071;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/454,557C
FILING DATE: 30-MAY-1995
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609,3840003
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2540
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 91:
SEQUENCE CHARACTERISTICS:
LENGTH: 84 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
US-08-454-557C-91

Query Match
Best Local Similarity 100.0%; Score 34; DB 3; Length 84;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2676 GCCCGCCTCAGCCTCCCAAGTCTGGATTACA 2709
DB 44 GCCCGCCTCAGCCTCCCAAGTCTGGATTACA 77

RESULT 14
US-08-340-426D-91
Sequence 91, Application US/08340426D
Patent No. 5948634
GENERAL INFORMATION:
APPLICANT: de la Monte, Suzanne
TITLE OF INVENTION: Neutral Thread Protein Gene Expression and Detection
NUMBER OF SEQUENCES: 121
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/340,426D
FILING DATE: 14-NOV-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609,3840002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 91:
SEQUENCE CHARACTERISTICS:
LENGTH: 84 base pairs
TYPE: nucleic acid

STRANDEDNESS: both
TOPOLOGY: both
US-08-340-426D-91

Query Match
Best Local Similarity 100.0%; Score 34; DB 4; Length 84;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2676 GCCCGCCTCAGCCTCCCAAGTCTGGATTACA 2709
DB 44 GCCCGCCTCAGCCTCCCAAGTCTGGATTACA 77

RESULT 15
US-08-450-673C-91
Sequence 91, Application US/08450673C
Patent No. 5948888
GENERAL INFORMATION:
APPLICANT: de la Monte, Suzanne
TITLE OF INVENTION: Neutral Thread Protein Gene Expression and Detection
NUMBER OF SEQUENCES: 121
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/450,673C
FILING DATE: 30-MAY-1995
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609,3840004
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 91:
SEQUENCE CHARACTERISTICS:
LENGTH: 84 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
US-08-450-673C-91

Query Match
Best Local Similarity 100.0%; Score 34; DB 4; Length 84;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2676 GCCCGCCTCAGCCTCCCAAGTCTGGATTACA 2709
DB 44 GCCCGCCTCAGCCTCCCAAGTCTGGATTACA 77

Search completed: March 30, 2000, 10:42:50
Job time: 111232 sec

Mon Apr 3 08:24:30 2000

us-08-852-495c-1_copy_65000_95000.rn1

Page 7

GenCore version 4.5
Copyright (c) 1993 - 2000 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 29, 2000, 20:38:58 ; Search time 6197.58 Seconds
(without alignments)
13817.943 Million cell updates/sec

Title: US-08-852-495c-1_COPY_65000_95000

Perfect score: 30001

Sequence: 1 TTATTATGACAAATTAAT.....TACAGATACTCTGCAGTTT 30001

Scoring table:

OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 4538634 seqs, 1887831982 residues

Word size : 0

Total number of hits satisfying chosen parameters: 130662

Minimum DB seq length: 8

Maximum DB seq length: 105

Post-processing: Listing first 45 summaries

Database : EST:*

1: em_est1:*
2: em_est2:*
3: em_est3:*
4: em_est4:*
5: em_est5:*
6: em_est6:*
7: em_est7:*
8: em_est8:*
9: em_est9:*
10: em_est10:*
11: em_est11:*
12: em_est12:*
13: em_est13:*
14: em_est14:*
15: em_est15:*
16: em_est16:*
17: em_est17:*
18: em_est18:*
19: em_est19:*
20: gb_est1:*
21: gb_est2:*
22: gb_est3:*
23: gb_est4:*
24: gb_est5:*
25: gb_est6:*
26: gb_est7:*
27: gb_est8:*
28: gb_est9:*
29: gb_est10:*
30: gb_est11:*
31: gb_est12:*
32: gb_est13:*
33: gb_est14:*
34: gb_est15:*
35: gb_est16:*
36: gb_est17:*
37: gb_est18:*
38: gb_est19:*
39: gb_est20:*
40: gb_est21:*
41: gb_est22:*
42: gb_est23:*
43: gb_est24:*

44: gb_est25:*
45: gb_est26:*
46: gb_est27:*
47: gb_est28:*
48: gb_est29:*
49: gb_est30:*
50: gb_est31:*
51: gb_est32:*
52: em_est20:*
53: em_est21:*
54: em_est22:*
55: em_est23:*
56: em_est24:*
57: em_est25:*
58: em_est26:*
59: gb_est33:*
60: gb_est34:*
61: gb_est35:*
62: gb_est36:*
63: gb_est37:*
64: gb_est38:*
65: em_est27:*
66: em_est28:*
67: em_est29:*
68: em_est30:*
69: em_est39:*
70: gb_est40:*
71: gb_est41:*
72: gb_est42:*
73: gb_est43:*
74: gb_est44:*
75: em_est31:*
76: em_est32:*
77: em_est33:*
78: em_est34:*
79: gb_est45:*
80: gb_est46:*
81: gb_est47:*
82: gb_est48:*
83: em_est34:*
84: em_est35:*
85: em_est36:*
86: em_est37:*
87: gb_est49:*
88: gb_est50:*
89: gb_est51:*
90: gb_est52:*
91: gb_est53:*
92: em_est38:*
93: em_est39:*
94: em_est40:*
95: em_est41:*
96: em_est42:*
97: em_est43:*
98: em_est44:*
99: gb_est54:*
100: gb_est55:*
101: em_est45:*
102: gb_est56:*
103: gb_est57:*
104: gb_est58:*
105: gb_est59:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	54	0.2	83	21	R07384	R07384.ye96c03.r1

RESULT	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1	R07384	ye96c03.r1 Soares fetal liver spleen INFIS Homo sapiens cDNA clone IMAGE:125572 5' similar to gb:X57130.cdsl HISTONE H1D (HUMAN);	R07384	83 bp	EST	05-APR-1995	Human	1 (bases 1 to 83)	Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevisan, E., Waterston, R., Williamson, A., Woldmann, P. and Wilson, R.	The WashU-Merck EST Project.	Unpublished (1995)	Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

ALIGNMENTS

RESULT	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
2	AA780764/c	ac68f12.s1 Striatogene fetal retina 937202 Homo sapiens cDNA clone IMAGE:867791 3' similar to contans Alu repetitive element; mRNA sequence.	AA780764.1	91 bp	EST	05-FEB-1998	Human	1 (bases 1 to 91)	Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S., Krizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J., Moore, B., Schellenberg, K., Stepien, M., Tan, F., Theising, B., White, Y., Wylie, T., Waterston, R. and Wilson, R.	Unpublished (1997)	On Sep 12, 1996 this sequence version replaced g1:1407381.	

Query Match 0.28; Score 54; DB 21; Length 83;
Best Local Similarity 100.0%; Pred. No. 1.6e-06;
Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13008 CGCGCTATGATGCGAGAAAACACACCGTCAACTGCTCTCAAGACCT 13061
Db 30 CGCGCTATGATGCGAGAAAACACACCGTCAACTGCTCTCAAGACCT 83

FEATURES
source
1. .83
/organism="Homo sapiens"
/db_xref="GDB:478117"
/db_xref="taxon:9606"
/clone="IMAGE:125572"
/clone_1bp="Soares fetal liver spleen INFIS"
/sex="male"
/dev_stage="20 week-post conception fetus"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: Liver and Spleen; Vector: pT73D (Pharmacia) with a modified polylinker; Site 1: Pac I; Site 2: Eco RI; 1st strand cDNA was primed with a Pac I - oligo(dT) primer [5', AACTGCAGCAATTAATTAAGCTCTTTTCTTTTCTTTTCTTTT 3', double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Pac I and cloned into the Pac I and Eco RI sites of the modified pT73 vector. Library went through one round of normalisation. Library constructed by Benito Soares and M. Fatima Bonaldo."

BASE COUNT
24 a 19 c 17 t

ORIGIN


```

RESULT 5
LOCUS AQ319270 102 bp DNA GSS 06-MAY-1999
DEFINITION RPI11-98B22.TV RPI1-11 Homo sapiens genomic clone RPI1-11-98B22,
ACCESSION AQ319270
VERSION AQ319270.1 GI:4052235
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 102)
AUTHORS Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.,
Berry,K., Granger,D., Suh,E., Wible,C., de Jong,P. and Venter,J.C.
TITLE Use of human BAC end sequences for Sequence-Ready Map Building
JOURNAL Unpublished (1998)
COMMENT Other GSSs: RPI11-98B22.TV
Contact: Shaying Zhao, William Nierman, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208
Email: hbeet@ig.org
Clones are derived from the human BAC library RPI1-11. For BAC
library availability, please contact Pieter de Jong
(pieter@dejong.med.bufralo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.bufralo.edu/ordering) or from
Research Genetics (info@resgen.com). BAC end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
Seq primer: SP6
Class: BAC ends.

FEATURES
source
location/Qualifiers
1..102
/organism="Homo sapiens"
/db_xref="GDB:7537293"
/db_xref="taxon:9606"
/clone="RPI1-11-98B22"
/clone_1lb="RPI1-11"
/sex="Male"
/cell_type="Lymphocytes"
/note="Vector: DBAC3.6; Site 1: EcoRI; Site 2: EcoRI;
RPI11 Human Male BAC Library"

BASE COUNT 26 a 28 c 31 g 17 t

Query Match 0.1%; Score 45; DB 100; Length 102;
Best Local Similarity 100.0%; Pred. No. 0.00084;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28844 GGTGATCACCCTCGCTCCCAAGTCTGGGATTACAGT 28888
Db 97 GGTGATCACCCTCGCTCCCAAGTCTGGGATTACAGT 53

RESULT 6
LOCUS F29274 91 bp mRNA EST 13-MAY-1999
DEFINITION HSPD19063 HM3 Homo sapiens cDNA clone s400074G04, mRNA sequence.
ACCESSION F29274
VERSION F29274.1 GI:4814900
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 91)
AUTHORS Lanfanchi,G., Murro,T., Caldara,F., Pacchioni,B., Pallavicini,A.,
Pandolfo,D., Toppo,S., Trevisan,S., Scarso,S. and Valle,G.

```

```

TITLE Identification of 4370 expressed sequence tags from a
3'-end-specific cDNA library of human skeletal muscle by DNA
sequencing and filter hybridization
JOURNAL Genome Res. 6 (1), 35-42 (1996)
MEDLINE 96276048
COMMENT On Apr 7, 1998 this sequence version replaced gi:3034419.
Contact: Valle G.
CIRIBI Biotechnology Centre
University of Padua
Via Trieste 75, 35121 Padua, Italy
ABI Chromatograms and other information are available on WWW at
http://grup.bio.unipd.it.

FEATURES
source
location/Qualifiers
1..91
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="s400074G04"
/clone_1lb="HM3"
/sex="female"
/tissue_type="pectoral muscle (after mastectomy)"
/note="Vector: pCDNAII (Invitrogen); Site 1: BstXI;
Site 2: NotI. This library is not subtracted nor normalized.
Lanfanchi. This library is not subtracted nor normalized.
The first strand cDNA was primed with a biotinylated
oligo-dT-NotI primer
ds cDNA was sonicated and size-selected in the range
350-550 bp. The 3' specific fragments were selected by
streptavidin coated magnetic beads, ligated to
non-palindromic BstXI adapters, NotI digested and
directionally cloned into BstXI-NotI cut pCDNAII vector."

BASE COUNT 18 a 19 c 38 g 16 t

Query Match 0.1%; Score 43; DB 50; Length 91;
Best Local Similarity 100.0%; Pred. No. 0.0037;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 26283 GGGCGCTGATGCCAGCTACGCGAGGCGTGGAGGAGAA 26325
Db 24 GGGCGCTGATGCCAGCTACGCGAGGCGTGGAGGAGAA 66

RESULT 7
LOCUS AU038446 93 bp mRNA EST 29-MAR-1999
DEFINITION AU038446 Dictyostelium discoideum SS (H.Urushihara) Dictyostelium
discoideum cDNA clone SSH/29, mRNA sequence.
ACCESSION AU038446
VERSION AU038446.1 GI:3985199
KEYWORDS EST.
SOURCE Dictyostelium discoideum.
ORGANISM Dictyostelium discoideum
Eukaryota; Dictyostelida; Dictyostelium.
REFERENCE 1 (bases 1 to 93)
AUTHORS Morio,T., Urushihara,H., Saito,T., Ugawa,Y., Mizuno,H., Yoshida,M.,
Yoshino,R., Mitra,B.N., Pi,M., Sato,T., Takemoto,K., Yasukawa,H.,
Williams,J., Maeda,M., Takeuchi,I., Ochial,H. and Tanaka,Y.
TITLE The Dictyostelium developmental cDNA project: generation and
analysis of expressed sequence tags from the first-finger stage of
development
JOURNAL DNA Res. 5 (6), 335-340 (1998)
MEDLINE 99156227
COMMENT On Dec 5, 1997 this sequence version replaced gi:2662913.
Contact: Hideko Urushihara
Institute of Biological Sciences
University of Tsukuba
3-3-10 Ten-noda, Tsukuba, Ibaraki 305, Japan
Email: d402huesakura.cc.tsukuba.ac.jp
PROJECT "Dictyostelium discoideum cDNA project in Japan".
location/Qualifiers
1..93

```

```

BASE COUNT      57 a      2 c      1 g      33 t
ORIGIN

Query Match      0.1% Score 43; DB 44; Length 93;
Best Local Similarity 100.0%; Pred. No. 0.0037;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0

OY   2146 AATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2188
       |||
Db    27 AATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 69
       |||

RESULT      8
LOCUS      AA457423              95 bp mRNA EST          06-JUN-1997
DEFINITION aa68b02.r1 Stratiogene fetal retina 937202 Homo sapiens cDNA clone
IMAG=838155 5' similar to contains Alu repetitive element;; mRNA
sequence.
ACCESSION   AA457423
VERSION     AA457423.1 GI:2180143
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE   1 (bases 1 to 95)
AUTHORS     Hillier,L., Allen,M., Bowles,L., Dubnue,T., Gelsel,G., Jost,S.,
            Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J.,
            Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B.,
            White,Y., Wyllie,T., Waterson,R. and Wilson,R.
            WashU-Merck EST Project 1997
JOURNAL     Unpublished (1997)
COMMENT     On Nov 4, 1996 this sequence version replaced gi:1671271.
            Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through ENL; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -28m13 rev1 ET from Amersham
            High quality sequence stop: 76.
FEATURES             source
     source           1..95
                        /organism="Homo sapiens"
                        /db_xref="taxon:9606"
                        /clone_image="838155"
                        /clone_id="Stratigene fetal retina 937202"
                        /sex="mixed"
                        /lab_host="SOLR (kanamycin resistant)"
                        /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
                        XhoI; cloned unidirectionally. Primer: Oligo dT. Pooled
                        retinal tissue. Average insert size: 1.0 kb; Uni-ZAP XR
                        Vector; ~5' adaptor sequence: 5' GAATCGGACGACAG 3' ~3'
                        adaptor sequence: 5' CTCACGCTTTT TTTT TTTT 3'"
BASE COUNT      14 a      39 c      25 g      17 t
ORIGIN

```

Db	84	TCCGAGCTACTCGGAGCCTGACGAGCAGGAATGCGGTGAC	43
RESULT	9		
LOCUS	AU072757	73 bp mRNA	EST 24-JUN-1999
DEFINITION	AU072757 Dictyostelium discoidum SS (H.Urushihara) Dictyostelium		
ACCESSION	AU072757	dictoideum cDNA clone SSR65, mRNA sequence.	
VERSION	AU072757.1	GI:5179178	
KEYWORDS	EST.		
SOURCE	Dictyostelium discoidum.		
ORGANISM	Eukaryota; Dictyostellida; Dictyostelium.		
REFERENCE	1 (bases 1 to 73)		
AUTHORS	Urushihara,H.		
TITLE	Developmental cDNA in Dictyostelium discoidum (1999)		
JOURNAL	Unpublished (1999)		
COMMENT	On Jun 5, 1998 this sequence version replaced gi:3188495. Contact: Hideko Urushihara Institute of Biological Sciences University of Tsukuba 3-3-10 Ten-noda, Tsukuba, Ibaraki 305, Japan Email: da02nuesakura.cc.tsukuba.ac.jp PROJECT - Dictyostelium discoidum cDNA project in Japan. Location/Qualifiers 1..73 /organism="Dictyostelium discoidum" /strain="AX4" /db_xref="taxon:44689" /clone="SSR65" /clone_id="Dictyostelium discoidum SS (H.Urushihara)" /dev_stage="slung"		
BASE COUNT	46 a 26 t		
ORIGIN			
Query Match	0.1%; Score 41; DB 51; Length 73;		
Best Local Similarity	100.0%; Pred. No. 0.017;		
Matches 41; Conservative	0; Mismatches 0; Indels 0; Gaps 0;		
QY	2146 AATAATATATATATATATATATATATATATATATATATATAA	2186	
Db	15 AATAATATATATATATATATATATATATATATATATATAA	55	
RESULT	10		
LOCUS	A1932967/c	92 bp mRNA	EST 02-SEP-1999
DEFINITION	wc0f09.x1 NCI_CGAP_Gas4 Homo sapiens CDNA clone IMAGE:2457833	3'	
ACCESSION	A1932967		
VERSION	A1932967.1	GI:5671704	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;		
AUTHORS	Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
TITLE	1 (bases 1 to 92)		
JOURNAL	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.		
COMMENT	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index Unpublished (1997) On Dec 20, 1995 this sequence version replaced gi:1134104. Contact: Robert Strausberg, Ph.D. Tel.: (301) 496-1550 Email: Robert_Strausberg@nih.gov Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D. CDNA Library Preparation: Life Technologies, Inc. CDNA Library Arrayed by: Greg Lennon, Ph.D. DNA Sequencing by: Washington University Genome Sequencing Center Clone distribution: NCI-CGAP clone distribution information can be		

Found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/dbrrp/image/image.html

Seq primer: -40up from Glpco.
Location/Qualifiers

FEATURES

SOURCE

1. 92
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="2457833"
/clone_lib="NCI-GAP-Gas4"
/tissue_type="Poorly differentiated adenocarcinoma with
signed ring cell features"
/lab_host="DH10B"
/note="Organ: stomach; Vector: pCMV-SPORT6; Site_1: Salt,
Site_2: Not; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.69 kb. Life technologies catalog #:
11549-011"

BASE COUNT 29 a 19 c 14 g 30 t
ORIGIN

Query Match 0.1%; Score 41; DB 62; Length 92;
Best Local Similarity 100.0%; Pred. No. 0.015;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 26392 GACAGACGAGACTCCGCTCAAAAAAAAAAAAAAAAA 26432

DB 41 GACAGACGAGACTCCGCTCAAAAAAAAAAAAAAAAA 1

RESULT 11
A0626158 104 bp DNA GSS 16-JUN-1999
LOCUS CTRBI-E1-2650P18.FR CTRBI-E1 Homo sapiens genomic clone 2650P18,
DEFINITION genomic survey sequence.
ACCESSION A0626158
VERSION A0626158.1 GI:5088550
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 104)
Zhao,S., Adams,M.D., Nierman,W., Malek,J., Shizuya,H., Simon,M. and
Venter,J.C.
Use of BAC End Sequences from Caltech Libraries for Sequence-Ready
Map Building
Unpublished (1997)
Other-GSS: CTRBI-E1-2650P18.FR
Contact: Shaying Zhao, William Nierman, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0208
Fax: 301 838 0208
Email: hbeetlgr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tdb/hungen/bac_end_search/bac_end_search.html.
Seq primer: M13 Reverse
Class: BAC ends.

JOURNAL
COMMENT
TITLE
AUTHORS
REFERENCE
A0626158 104 bp DNA GSS 16-JUN-1999
LOCUS CTRBI-E1-2650P18.FR CTRBI-E1 Homo sapiens genomic clone 2650P18,
DEFINITION genomic survey sequence.
ACCESSION A0626158
VERSION A0626158.1 GI:5088550
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 104)
Zhao,S., Adams,M.D., Nierman,W., Malek,J., Shizuya,H., Simon,M. and
Venter,J.C.
Use of BAC End Sequences from Caltech Libraries for Sequence-Ready
Map Building
Unpublished (1997)
Other-GSS: CTRBI-E1-2650P18.FR
Contact: Shaying Zhao, William Nierman, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0208
Fax: 301 838 0208
Email: hbeetlgr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tdb/hungen/bac_end_search/bac_end_search.html.
Seq primer: M13 Reverse
Class: BAC ends.

FEATURES
SOURCE
Location/Qualifiers

1. 104
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="2650P18"
/clone_lib="CTRBI-E1"
/sex="male"
/cell_type="sperm"
/note="Vector: pBeloBAC11; Site_1: EcoRI; Site_2: EcoRI;
Caltech Human BAC Library D"

BASE COUNT 12 a 28 c 22 g 42 t
ORIGIN

Query Match 0.1%; Score 41; DB 105; Length 104;
Best Local Similarity 100.0%; Pred. No. 0.014;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 26392 GACAGACGAGACTCCGCTCAAAAAAAAAAAAAAAAA 26432

DB 94 GACAGACGAGACTCCGCTCAAAAAAAAAAAAAAAAA 54

RESULT 12
D25879 52 bp mRNA EST 30-NOV-1995
LOCUS HMG505672 Human colon mucosa Homo sapiens CDNA clone cm2333 3',
DEFINITION mRNA sequence.
ACCESSION D25879
VERSION D25879.1 GI:500543
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 52)
Okubo,K., Yoshii,J., Yokouchi,H., Kameyama,M. and Matsubara,K.
Global analysis of gene expression in colon mucosa: a large scale
random CDNA sequencing analysis
Unpublished (1994)
Contact: Okubo,K., Itoh,K., Yoshii,J., Yokouchi,H. and Matsubara,K.
Institute for Molecular and Cellular Biology
Osaka University
5-1 Yamadaoka, Suita, Osaka 565, Japan.
Location/Qualifiers

FEATURES
SOURCE
Location/Qualifiers

1. 52
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="cm2333"
/clone_lib="Human colon mucosa"
/note="Adult male, tissue_type = colon mucosa"

BASE COUNT 12 a 17 c 14 g 9 t
ORIGIN

Query Match 0.1%; Score 40; DB 20; Length 52;
Best Local Similarity 100.0%; Pred. No. 0.043;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 26376 GCACTCCAGCGCTGGTGACAGAGCGAGACTCCGCTCAA 26415

DB 13 GCACTCCAGCGCTGGTGACAGAGCGAGACTCCGCTCAA 52

RESULT 13
AA907458 97 bp mRNA EST 26-AUG-1998
LOCUS o102h06.s1 NCI-CGAP_Lus Homo sapiens CDNA clone IMAGE:1522331 3',
DEFINITION similar to contigAlu repetitive element; mRNA sequence.
ACCESSION AA907458
VERSION AA907458.1 GI:3042918
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 97)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
On Jan 14, 1998 this sequence version replaced gi:1877755.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert_Strausberg@nih.gov

JOURNAL
COMMENT
TITLE
AUTHORS
REFERENCE
AA907458 97 bp mRNA EST 26-AUG-1998
LOCUS o102h06.s1 NCI-CGAP_Lus Homo sapiens CDNA clone IMAGE:1522331 3',
DEFINITION similar to contigAlu repetitive element; mRNA sequence.
ACCESSION AA907458
VERSION AA907458.1 GI:3042918
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 97)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
On Jan 14, 1998 this sequence version replaced gi:1877755.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert_Strausberg@nih.gov

BASE COUNT 12 a 28 c 22 g 42 t
ORIGIN

Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
 CDNA Library Preparation: M. Bento Soares, Ph.D.
 CDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www.bio.llnl.gov/dbp/image/image.html

Insert Length: 624 Std Error: 0.00
 Seq primer: -40ml3 fwd. ET from Amersham
 High quality sequence stop: 97.
 Location/Qualifiers

FEATURES

source

1..97
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_image="1522331"
 /clone_lib="NCI-CGAP.Lu5"
 /issue_type="carcinoid"
 /lab_host="DH108"
 /note="Organ: lung; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; 1st strand cDNA was prepared from neuroendocrine lung carcinoma, and was then primed with a Not I - oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library is normalized. Library was constructed by Bento Soares and M. Fatima Bonaldo."
 BASE COUNT 17 a 28 c 23 g 27 t 2 others
 ORIGIN

Query Match 0.1%; Score 40; DB 39; Length 97;
 Best Local Similarity 100.0%; Pred. No. 0.03;
 Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8636 TTGTACCCAGGCTGAGTGGCGCAATCTCAGCTCA 8675
 B48914
 LOCUS 5 TTGTACCCAGGCTGAGTGGCGCAATCTCAGCTCA 44

RESULT 14
 B48914 103 bp DNA GSS 08-APR-1999
 LOCUS RPII11-4A12.TP RPII-11 Homo sapiens genomic clone RPII-11-4A12,
 DEFINITION genomic survey sequence.
 B48914
 ACCESSION B48914.1 GI:2601151
 VERSION
 KEYWORDS GSS.
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 103)
 AUTHORS Adams,M.D., Rounsley,S.D., Field,C.E., Bass,S., Linher,K., Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., de Jong,P. and Venter,U.C.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

USE OF BAC END SEQUENCES FOR SEQUENCE-READY MAP BUILDING
 UNPUBLISHED (1997)
 CONTACT: Mark Adams
 DEPARTMENT OF Eukaryotic Genomics
 THE INSTITUTE FOR GENOMIC RESEARCH
 9712 Medical Center Dr., Rockville, MD 20850, USA
 TEL: 301 838 0200
 FAX: 301 838 0208

EMAIL: mdadams@tigr.org
 CLONES ARE DERIVED FROM THE HUMAN BAC LIBRARY RPII-11. FOR BAC LIBRARY AVAILABILITY, PLEASE CONTACT PIETER DE JONG (pieter@ejeon.med.buffalo.edu). CLONES MAY BE PURCHASED FROM BACPAC RESOURCES (<http://bacpac.med.buffalo.edu/ordering>) OR FROM RESEARCH GENETICS (info@resgen.com). BAC END SEARCH PAGE: http://www.tigr.org/tldb/humgen/bac_end_search/bac_end_search.html
 Seq primer: Sp6

Class: BAC ends.
 Location/Qualifiers

FEATURES

source

1..103
 /organism="Homo sapiens"
 /db_xref="GDB:7501163"
 /db_xref="taxon:9606"
 /clone_image="RPII-11-4A12"
 /clone_lib="RPII-11"
 /sex="Male"
 /cell_type="Lymphocytes"
 /note="Vector: pBAC3.6; Site 1: EcoRI; Site 2: EcoRI; RPII11 Human Male BAC Library"
 BASE COUNT 30 a 28 c 30 g 15 t
 ORIGIN

Query Match 0.1%; Score 40; DB 81; Length 103;
 Best Local Similarity 100.0%; Pred. No. 0.029;
 Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 26168 TTGTGGAGCGCCGAGCGCGGATCAGAGCTCAGAGAT 26207
 B48914
 LOCUS 14 TTGTGGAGCGCCGAGCGCGGATCAGAGCTCAGAGAT 53

RESULT 15
 C25772 58 bp mRNA EST 29-JUL-1997
 LOCUS C25772 Dictyostelium discoideum FC (H. Urushihara) Dictyostelium
 DEFINITION dictyosideum cDNA clone FC-BB11, mRNA sequence.
 C25772
 ACCESSION C25772.1 GI:2282558
 VERSION
 KEYWORDS EST.
 SOURCE Dictyostelium discoideum.
 ORGANISM Dictyostelium discoideum.

REFERENCE 1 (bases 1 to 58)
 AUTHORS Suzuki,K., Shimizu,H. and Urushihara,H.
 TITLE Sexual cDNA in D. discoideum(970724)
 JOURNAL Unpublished (1997)
 COMMENT On Apr 14, 1993 this sequence version replaced g1:693374.
 CONTACT: Hideko Urushihara
 INSTITUTE OF BIOLOGICAL SCIENCES
 UNIVERSITY OF TSUKUBA
 3-3-10 Ten-nodai, Tsukuba, Ibaraki 305, Japan
 EMAIL: d402nuesakura.cc.tsukuba.ac.jp.

FEATURES
 source
 1..58
 /organism="Dictyostelium discoideum"
 /strain="KAX3"
 /db_xref="taxon:44689"
 /clone_image="FC-BB11"
 /clone_lib="Dictyostelium discoideum FC (H. Urushihara)"

BASE COUNT 35 a 6 c 1 g 16 t
 ORIGIN

Query Match 0.1%; Score 39; DB 35; Length 58;
 Best Local Similarity 100.0%; Pred. No. 0.082;
 Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2147 ATATATATATATATATATATATATATATATATATA 2165
 B48914
 LOCUS 1 ATATATATATATATATATATATATATATATATATA 39

Search completed: March 30, 2000, 03:49:55
 Job time: 88449 sec

Mon Apr 3 08:24:30 2000

us-08-852-495c-1_copy_65000_95000.rst

Page 8

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 30, 2000, 07:33:31 ; Search time 13412.7 Seconds

(without alignments) updates/sec

Title: US-08-852-495c-1_COPY_90000_120000

Sequence: 1 TGGCCAGCTGCTGAGTGGCGG.....TATTCATAGCCATGTCAT 30001

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 821193 seqs, -1518192014 residues

Word size : 0

Total number of hits satisfying chosen parameters: 345848

Minimum DB seq length: 8

Maximum DB seq length: 105

Post-processing: Listing first 45 summaries

Database :

GenBank: *
1: gb_ba1: *
2: gb_ba2: *
3: gb_com: *
4: gb_ov: *
5: gb_pat: *
6: gb_ph: *
7: gb_pl1: *
8: gb_pl2: *
9: gb_pl3: *
10: gb_pr2: *
11: gb_pr3: *
12: gb_ro: *
13: gb_sts: *
14: gb_sy: *
15: gb_un: *
16: gb_v1: *
17: em_fun: *
18: em_hum1: *
19: em_hum2: *
20: em_hum3: *
21: em_com: *
22: em_or: *
23: em_ov: *
24: em_pat: *
25: em_ph: *
26: em_pl: *
27: em_ro: *
28: em_sts: *
29: em_sy: *
30: em_un: *
31: em_v1: *
32: gb_htg1: *
33: gb_htg2: *
34: gb_htg3: *
35: gb_htg4: *
36: em_ba1: *
37: em_ba2: *
38: em_hum4: *
39: em_hum5: *
40: gb_pr4: *
41: gb_htg5: *
42: gb_htg6: *
43: gb_htg7: *

44: gb_htg8: *
45: em_htg1: *
46: em_htg2: *
47: em_htg3: *
48: em_htg4: *
49: em_htg5: *
50: gb_pl3: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	58	0.2	76	5 AR051499	AR051499 Sequence
2	55	0.2	60	5 AR051487	AR051487 Sequence
3	48	0.2	48	5 AR032547	AR032547 Sequence
4	48	0.2	48	5 I29287	I29287 Sequence 15
5	48	0.2	48	5 I29287	I29287 Sequence 15
6	47	0.2	47	5 AR032558	AR032558 Sequence 15
7	47	0.2	47	5 I29298	I29298 Sequence 17
8	47	0.2	47	5 I29298	I29298 Sequence 17
9	47	0.2	47	5 AR032555	AR032555 Sequence 17
10	46	0.2	46	5 I29285	I29285 Sequence 16
11	46	0.2	46	5 I29285	I29285 Sequence 16
12	45	0.1	45	5 AR032557	AR032557 Sequence 16
13	45	0.1	45	5 I29297	I29297 Sequence 16
14	45	0.1	45	5 I29297	I29297 Sequence 16
15	37	0.1	79	10 S73203	S73203 ATL-1 (rand
16	36	0.1	66	40 AF087511	AF087511 Homo sapi
17	35	0.1	57	13 HUMOT1543A	HUMOT1543A Homo sap
18	34	0.1	42	9 HUMALUANC	HUMALUANC Homo sapien
19	33	0.1	65	9 HUMALUANC	HUMALUANC Homo sapien
20	32	0.1	63	3 DOGSR7SL1	DOGSR7SL1 Dog signa
21	32	0.1	80	9 HUMBRKFC	HUMBRKFC Human alpha
22	31	0.1	66	10 S54071	S54071 hypoxanthin
23	31	0.1	67	13 HUMOT44A	HUMOT44A Human STS U
24	31	0.1	69	13 HUMOT578A	HUMOT578A Human STS U
25	31	0.1	90	9 HUMDLRFL	HUMDLRFL Human low d
26	31	0.1	100	11 HS067848	HS067848 Human beta-
27	31	0.1	101	13 HUMOT832A	HUMOT832A Human STS U
28	31	0.1	104	5 A08899	A08899 H. sapiens (
29	30	0.1	30	5 A38240	A38240 Sequence 4
30	30	0.1	30	5 A38241	A38241 Sequence 5
31	30	0.1	30	5 A38242	A38242 Sequence 6
32	30	0.1	66	40 AF087511	AF087511 Homo sapi
33	30	0.1	67	13 HUMOT44A	HUMOT44A Homo sapi
34	30	0.1	73	5 A08915	A08915 H. sapiens (
35	30	0.1	93	9 HSSTHPKIB	HSSTHPKIB H. sapiens m
36	30	0.1	100	9 HSLSA27	HSLSA27 H. sapiens D
37	30	0.1	102	5 A08911	A08911 H. sapiens (
38	30	0.1	105	13 HS015Xa3	HS015Xa3 H. sapiens (
39	29	0.1	40	5 A68621	A68621 Sequence 1
40	29	0.1	65	9 HUMALUANC	HUMALUANC Homo sapien
41	29	0.1	75	11 AF032333	AF032333 Propithec
42	29	0.1	76	11 AF032287	AF032287 Eulemur m
43	29	0.1	88	13 HUMOT5056B	HUMOT5056B Human STS U
44	29	0.1	99	5 A08898	A08898 H. sapiens (
45	29	0.1	100	9 HUMGALNSA	HUMGALNSA Human GALNS

ALIGNMENTS

RESULT 1
LOCUS AR051499 76 bp DNA
DEFINITION AR051499 Sequence 69 from patent US 5830670.
ACCESSION AR051499
VERSION AR051499.1 GI:5974863

29-SEP-1999

Mon Apr 3 08:24:31 2000

us-08-852-495c-1_copy_90000_120000.1rge

Page 2

KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 76)
TITLE de la Monte,S. and Wands,J.R.
JOURNAL Neutral thread protein gene expression and detection of Alzheimer's
FEATURES disease
Patent: US 5830670-A 69 03-NOV-1998;
LOCATION/Qualifiers
SOURCE 1..76
BASE COUNT 15 a 19 c 18 g 24 t
ORIGIN

Query Match 0.2%; Score 58; DB 5; Length 76;
Best Local Similarity 100.0%; Pred. No. 1.1e-18;
Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 2685 GCGCCGCTAATTTTGTATTTTGTAGTACAGACAGGCTTCACCGTGTGGCCAGGATG 2742
Db 6 GCGCCGCTAATTTTGTATTTTGTAGTACAGACAGGCTTCACCGTGTGGCCAGGATG 63

RESULT 2
LOCUS AR051487 60 bp DNA PAT 29-SEP-1999
DEFINITION Sequence 57 from patent US 5830670.
ACCESSION AR051487
VERSION AR051487.1 GI:5974851
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 60)
AUTHORS de la Monte,S. and Wands,J.R.
TITLE Neutral thread protein gene expression and detection of Alzheimer's
JOURNAL disease
Patent: US 5830670-A 57 03-NOV-1998;
FEATURES Location/Qualifiers
SOURCE 1..60
BASE COUNT 12 a 14 c 15 g 19 t
ORIGIN

Query Match 0.2%; Score 55; DB 5; Length 60;
Best Local Similarity 100.0%; Pred. No. 4.1e-17;
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 2685 GCGCCGCTAATTTTGTATTTTGTAGTACAGACAGGCTTCACCGTGTGGCCAGG 2739
Db 6 GCGCCGCTAATTTTGTATTTTGTAGTACAGACAGGCTTCACCGTGTGGCCAGG 60

RESULT 3
LOCUS AR032547 48 bp DNA PAT 29-SEP-1999
DEFINITION Sequence 159 from patent US 5869241.
ACCESSION AR032547
VERSION AR032547.1 GI:5948152
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 48)
AUTHORS Edwards,C.A., Cantor,C.R., Andrews,B.M., Turin,L.M. and Fry,K.E.
TITLE Method of determining DNA sequence preference of a DNA-binding
JOURNAL molecule
Patent: US 5869241-A 159 09-FEB-1999;
FEATURES Location/Qualifiers
SOURCE 1..48

BASE COUNT 12 a 8 c 14 g 14 t
ORIGIN

Query Match 0.2%; Score 48; DB 5; Length 48;
Best Local Similarity 100.0%; Pred. No. 2e-13;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 2282 GAGAGAAATGGAACCGATCTTATATATCTACGTTACCCCTGCCCC 22629
Db 48 GAGAGAAATGGAACCGATCTTATATATCTACGTTACCCCTGCCCC 1

RESULT 4
LOCUS I29287 48 bp DNA PAT 06-FEB-1997
DEFINITION Sequence 159 from patent US 5578444.
ACCESSION I29287
VERSION I29287.1 GI:1820078
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 48)
AUTHORS Edwards,C.A., Cantor,C.R., Andrews,B.M., Turin,L.M. and Fry,K.E.
TITLE Sequence-directed DNA-binding molecules compositions and methods
JOURNAL Patent: US 5578444-A 159 26-NOV-1996;
FEATURES Location/Qualifiers
SOURCE 1..48
BASE COUNT 12 a 8 c 14 g 14 t
ORIGIN

Query Match 0.2%; Score 48; DB 5; Length 48;
Best Local Similarity 100.0%; Pred. No. 2e-13;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 2282 GAGAGAAATGGAACCGATCTTATATATCTACGTTACCCCTGCCCC 22629
Db 48 GAGAGAAATGGAACCGATCTTATATATCTACGTTACCCCTGCCCC 1

RESULT 5
LOCUS I90961 48 bp DNA PAT 01-DEC-1998
DEFINITION Sequence 159 from patent US 5726014.
ACCESSION I90961
VERSION I90961.1 GI:3935431
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 48)
AUTHORS Edwards,C.A., Cantor,C.R., Andrews,B.M. and Turin,L.M.
TITLE Screening assay for the detection of DNA-binding molecules
JOURNAL Patent: US 5726014-A 159 10-MAR-1998;
FEATURES Location/Qualifiers
SOURCE 1..48
BASE COUNT 12 a 8 c 14 g 14 t
ORIGIN

[illegible]

AUTHORS	Edwards,C.A., Cantor,C.R., Andrews,B.M. and Turlin,L.M.
TITLE	Screening assay for the detection of DNA-binding molecules
JOURNAL	Patent: US 5726014-A 170 10-MAR-1998;
FEATURES	Location/Qualifiers source 1..47
BASE COUNT	/organism="unknown" ORIGIN 13 a 8 c 10 g 16 t
Query Match	Best Local Similarity 0.2%; Score 47; DB 5; Length 47; Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 12078	CATAGTGGTGCTGACTATAAAGAAGTAGCTTTCCCT 12124 CATAGTGGTGCTGACTATAAAGAAGTAGCTTTCCCT 47
RESULT 9	AR032555 46 bp DNA PAT 29-SEP-1999
LOCUS	AR032555 46 bp DNA
DEFINITION	Sequence 167 from patent US 5869241.
ACCESSION	AR032555
VERSION	AR032555.1 GI:5948160
KEYWORDS	
SOURCE	Unknown. ORGANISM Unknown. UNCLASSIFIED.
REFERENCE	1 (bases 1 to 46) AUTHORS Edwards,C.A., Cantor,C.R., Andrews,B.M., Turlin,L.M. and Fry,K.E.
TITLE	Method of determining DNA sequence preference of a DNA-binding molecule
JOURNAL	Patent: US 5869241-A 167 09-FEB-1999;
FEATURES	Location/Qualifiers source 1..46
BASE COUNT	/organism="unknown" ORIGIN 10 a 16 c 6 g 14 t
Query Match	Best Local Similarity 0.2%; Score 46; DB 5; Length 46; Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 10562	CACAGCCTACCTCCAGTCAGTAAATACCTCTCGCTGGCTTC 10607 CACAGCCTACCTCCAGTCAGTAAATACCTCTCGCTGGCTTC 46
Db 1	CACAGCCTACCTCCAGTCAGTAAATACCTCTCGCTGGCTTC 46
RESULT 10	I29295 46 bp DNA PAT 06-FEB-1997
LOCUS	I29295 46 bp DNA
DEFINITION	Sequence 167 from patent US 5578444.
ACCESSION	I29295
VERSION	I29295.1 GI:1820086
KEYWORDS	
SOURCE	Unknown. ORGANISM Unknown. UNCLASSIFIED.
REFERENCE	1 (bases 1 to 46) AUTHORS Edwards,C.A., Cantor,C.R., Andrews,B.M., Turlin,L.M. and Fry,K.E.
TITLE	Sequence-directed DNA-binding molecules compositions and methods
JOURNAL	Patent: US 5578444-A 167 26-NOV-1996;
FEATURES	Location/Qualifiers source 1..46
BASE COUNT	/organism="unknown" ORIGIN 10 a 16 c 6 g 14 t
Query Match	Best Local Similarity 0.2%; Score 46; DB 5; Length 46; Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10562 CACAGCCTACCTCCAGTCAGTAAATAGTCTCTCGCTGCGTTC 10607
|||||
DB 1 CACAGCCTACCTCCAGTCAGTAAATAGTCTCTCTCGCTGCGTTC 46

RESULT 11
LOCUS 190969 46 bp DNA PAT 01-DEC-1998
DEFINITION Sequence 167 from patent US 5726014.
ACCESSION 190969
VERSION 190969.1 GI:3935439
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 46)
AUTHORS Edwards, C.A., Cantor, C.R., Andrews, B.M. and Turin, L.M.
TITLE Screening assay for the detection of DNA-binding molecules
JOURNAL Patent: US 5726014-A 167 10-MAR-1998;
FEATURES
Location/Qualifiers
1..46
/organism="unknown"
BASE COUNT 10 a 16 c 6 g 14 t
ORIGIN

Query Match 0.2%; Score 45; DB 5; Length 46;
Best Local Similarity 100.0%; Pred. No. 2.4e-12;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10562 CACAGCCTACCTCCAGTCAGTAAATAGTCTCTCGCTGCGTTC 10607
|||||
DB 1 CACAGCCTACCTCCAGTCAGTAAATAGTCTCTCTCGCTGCGTTC 46

RESULT 12
LOCUS AR032557 45 bp DNA PAT 29-SEP-1999
DEFINITION Sequence 169 from patent US 5669241.
ACCESSION AR032557
VERSION AR032557.1 GI:5948162
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 45)
AUTHORS Edwards, C.A., Cantor, C.R., Andrews, B.M., Turin, L.M. and Fry, K.E.
TITLE Method of determining DNA sequence preference of a DNA-binding molecule
JOURNAL Patent: US 5669241-A 169 09-FEB-1999;
FEATURES
Location/Qualifiers
1..45
/organism="unknown"
BASE COUNT 8 a 6 c 10 g 21 t
ORIGIN

Query Match 0.1%; Score 45; DB 5; Length 45;
Best Local Similarity 100.0%; Pred. No. 7.4e-12;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12057 ATTTTGAATTTCTTGCGTCCAAATAGTGTGCTGACTCTAT 12101
|||||
DB 1 ATTTTGAATTTCTTGCGTCCAAATAGTGTGCTGACTCTAT 45

RESULT 13
LOCUS 129297 45 bp DNA PAT 06-FEB-1997
DEFINITION Sequence 169 from patent US 5578444.
ACCESSION 129297

VERSION 129297.1 GI:1820088
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 45)
AUTHORS Edwards, C.A., Cantor, C.R., Andrews, B.M., Turin, L.M. and Fry, K.E.
TITLE Sequence directed DNA-binding molecules compositions and methods
JOURNAL Patent: US 5578444-A 169 26-NOV-1996;
FEATURES
Location/Qualifiers
1..45
/organism="unknown"
BASE COUNT 8 a 6 c 10 g 21 t
ORIGIN

Query Match 0.1%; Score 45; DB 5; Length 45;
Best Local Similarity 100.0%; Pred. No. 7.4e-12;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12057 ATTTTGAATTTCTTGCGTCCAAATAGTGTGCTGACTCTAT 12101
|||||
DB 1 ATTTTGAATTTCTTGCGTCCAAATAGTGTGCTGACTCTAT 45

RESULT 14
LOCUS 190971 45 bp DNA PAT 01-DEC-1998
DEFINITION Sequence 169 from patent US 5726014.
ACCESSION 190971
VERSION 190971.1 GI:3935441
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 45)
AUTHORS Edwards, C.A., Cantor, C.R., Andrews, B.M. and Turin, L.M.
TITLE Screening assay for the detection of DNA-binding molecules
JOURNAL Patent: US 5726014-A 169 10-MAR-1998;
FEATURES
Location/Qualifiers
1..45
/organism="unknown"
BASE COUNT 8 a 6 c 10 g 21 t
ORIGIN

Query Match 0.1%; Score 45; DB 5; Length 45;
Best Local Similarity 100.0%; Pred. No. 7.4e-12;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12057 ATTTTGAATTTCTTGCGTCCAAATAGTGTGCTGACTCTAT 12101
|||||
DB 1 ATTTTGAATTTCTTGCGTCCAAATAGTGTGCTGACTCTAT 45

RESULT 15
LOCUS S73203/c 79 bp DNA PRI 28-FEB-1995
DEFINITION ALU-1 [tandem duplication] [human, acute myeloid leukemia patient,
Genomic Mutant, 79 nt].
ACCESSION S73203
VERSION S73203.1 GI:685048
KEYWORDS
SOURCE human acute myeloid leukemia patient.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS Schichman, S.A., Caligiuri, M.A., Strout, M.P., Carter, S.L., Gu, Y.,
Canaani, B., Bloomfield, C.D. and Croce, C.M.
TITLE ALU-1 tandem duplication in acute myeloid leukemia with a normal
karyotype involves homologous recombination between Alu elements
JOURNAL Cancer Res. 54 (16), 4277-4280 (1994)

MEDLINE 94320053
 Genbank staff at the National Library of Medicine created this
 entry [NCBI g1dbsq 155288] from the original journal article.

This sequence comes from Fig. 3a.
 Map location: 11.

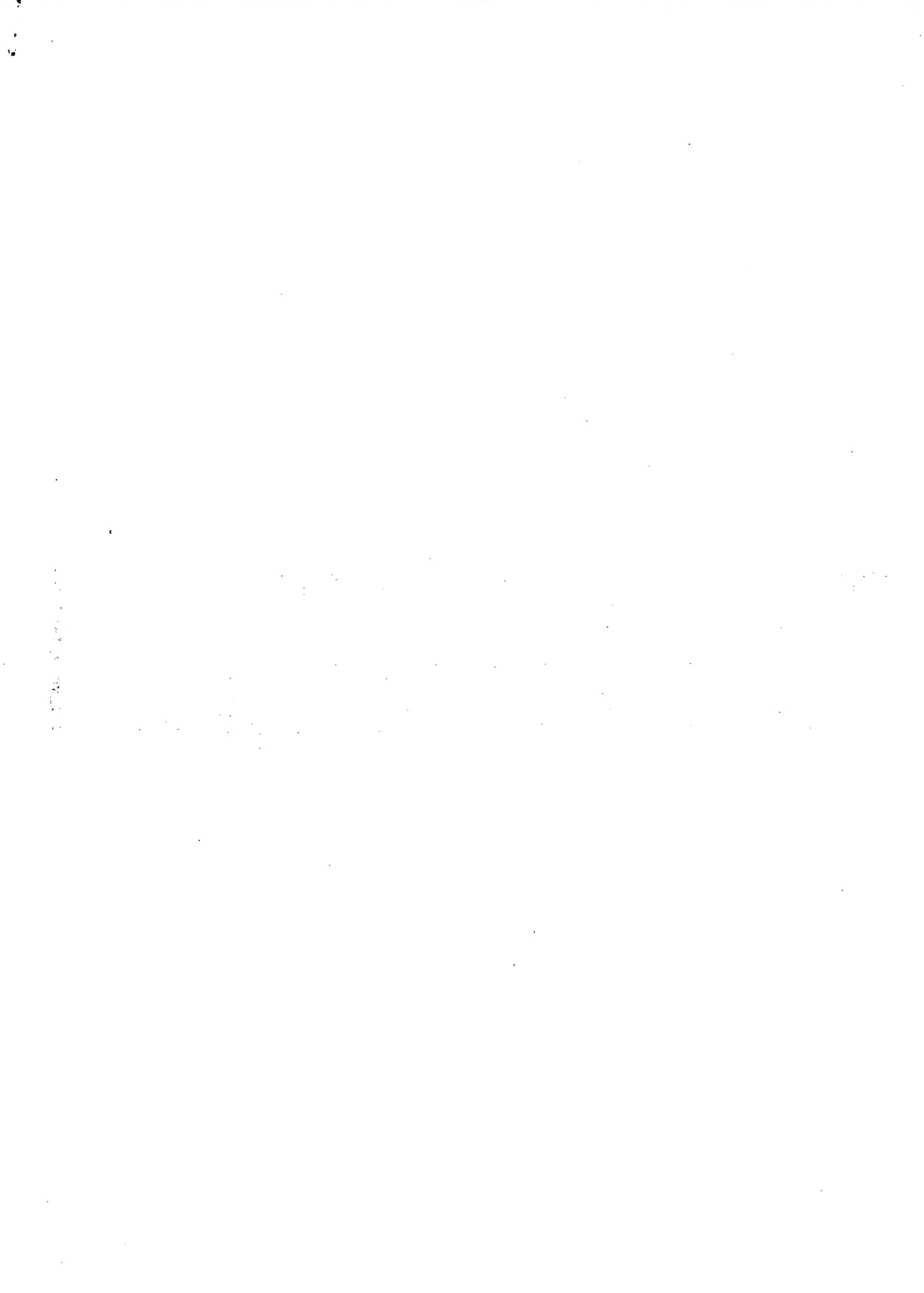
FEATURES
 source

1..79
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 1..79
 /partial
 /gene="ALT-1"

BASE COUNT 19 a 18 c 29 g 13 t
 ORIGIN

Query Match 0.1%; Score 37; DB 10; Length 79;
 Best Local Similarity 100.0%; Pred. No. 1.2e-07;
 Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 5166 TCTCAGCTCACTGCACCTCGCCTCCTGGSTCAAG 5202
 ||||||||||||||||||||||||||||||||||||
 Db 74 TCTCAGCTCACTGCACCTCGCCTCCTGGSTCAAG 38

Search completed: March 30, 2000, 17:03:40
 Job time: 135569 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 30, 2000, 10:53:10 ; Search time 611.3 Seconds
(without alignments)
12278.771 Million cell updates/sec

Title: US-08-852-495c-1_COPY_90000_120000
Perfect score: 30001
Sequence: 1 TGGCCAGCTGCGAGTGGCGG.....TATCTAGCAGTGTCAAT 30001

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 311585 seqs, 125096042 residues

Word size : 0

Total number of hits satisfying chosen parameters: 433264

Minimum DB seq length: 8
Maximum DB seq length: 105

Post-processing: Listing first 45 summaries

Database : N_Geneseq_36:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	48	0.2	48	069409	Human H4/a gene fo
2	48	0.2	48	T63871	Human H4/a gene (f
3	48	0.2	48	T17159	Test sequence from
4	47	0.2	47	069420	Human H3 histone g
5	47	0.2	47	T63882	Human H3 histone g
6	47	0.2	47	T17170	Test sequence from
7	46	0.2	46	069417	Human H2a histone
8	46	0.2	46	T63879	Human H2a histone
9	46	0.2	46	T17167	Test sequence from
10	45	0.1	45	069419	Human H3 histone g
11	45	0.1	45	T63881	Human H3 histone g
12	45	0.1	45	T17169	Test sequence from
13	40	0.1	30	062216	Histone H3 mRNA pr
14	30	0.1	30	062217	Histone H3 mRNA pr
15	30	0.1	30	062218	Histone H3 mRNA pr
16	29	0.1	40	T19044	Alu PCR primer 1.
17	28	0.1	60	X19944	OLIGONUCLEOTIDE SE
18	28	0.1	63	X19944	Microsatellite seq
19	28	0.1	68	T23904	Human gene signatu
20	27	0.1	36	V09783	Tobacco PABF probe
21	27	0.1	44	033855	Microsatellite seq
22	27	0.1	44	033855	Microsatellite seq
23	27	0.1	57	T65739	Repeat sequence fr
24	27	0.1	62	034053	Microsatellite seq
25	27	0.1	70	034155	Sequence of a micr
26	26	0.1	32	027389	Inter-Alu specific
27	26	0.1	35	027391	Inter-Alu specific
28	26	0.1	36	V09783	Tobacco PABF probe
29	26	0.1	42	086155	Sindbis polyA prim
30	26	0.1	42	T30789	Sindbis genomic cd
31	26	0.1	42	T35054	Sindbis virus stra
32	26	0.1	42	V42366	PCR primer 4B used
33	26	0.1	42	V60127	PCR primer 4B used

34	26	0.1	42	1	V70686	Primer 4B used to
35	26	0.1	48	1	086183	Primer SINdball1700
36	26	0.1	48	1	086183	Primer SINdball1700
37	26	0.1	48	1	T30807	Sindbis PCR primer
38	26	0.1	48	1	T30807	Sindbis PCR primer
39	26	0.1	48	1	T35073	Sindbis-based, tum
40	26	0.1	48	1	T35073	Sindbis-based, tum
41	26	0.1	48	1	V42384	Reverse PCR primer
42	26	0.1	48	1	V42384	Reverse PCR primer
43	26	0.1	48	1	V70704	Reverse PCR primer
44	26	0.1	48	1	V70704	Reverse PCR primer
45	26	0.1	53	1	033621	Microsatellite seq

ALIGNMENTS

```
RESULT 1
ID 069409/c
AC 069409;
DT 27-FEB-1995 (first entry)
DE Human H4/a gene for H4 histone, target region.
KW DNA protein-binding assay; test sequence; screening sequence;
KW promoter; target; TATA box; Herpes Simplex Virus; HSV;
KW origin of replication; UL5; transcription factor; TRID; ds.
OS Synthetic.
PN W09414980-A.
PD 07-JUL-1994.
PE 20-DEC-1993; U12388.
PR 23-DEC-1992; US-996783.
PR 17-SEP-1993; US-123936.
PA (GENE-) GENELABS TECHNOLOGIES INC.
PI Andrews BM, Cantor CR, Edwards CA, Fry KE, Turin LM;
DR WPI: 94-234711/28.
PT Sequence-directed DNA-binding molecules - useful in
PT pharmaceuticals and as molecular reagents
PS Claim 28: Page 291: 58pp; English.
CC A DNA protein-binding assay is provided, useful for screening
CC libraries of synthetic or biological cpds. for their ability
CC to bind DNA test sequences. The assay is versatile in that any
CC number of test sequences can be tested by placing the test sequence
CC adjacent to a defined protein-binding screening sequence. Binding
CC of mols. to these test sequences changes the binding characteristics
CC of the protein mol. to its cognate binding sequence. When such a mol.
CC binds the test sequence, the equilibrium of the DNA:protein complexes
CC is disturbed, generating changes in the concentration of free DNA probe.
CC One application of this method is to eucaryotic general transcription
CC factors (e.g. TFIID), where the target region is typically selected
CC from DNA sequences adjacent to the binding site for the eucaryotic
CC transcription factor. Numerous exemplary test sequences are given:
CC the sequences in Q69251-731 and Q69550 correspond to promoter targets
CC (typically, TATA box-contg. sites) for human genes and the sequences in
CC Q69733-849 correspond to promoter targets for viral genes. The test
CC sequences may also be randomly generated. DNA:protein interaction may
CC be used for screening purposes, e.g. the Herpes Simplex Virus (HSV)
CC origin of replication and UL5 (see Q69551-52, Q69865 and Q69891).
SQ Sequence 48 BP; 12 A; 8 C; 14 G; 14 T;
```

Query Match 0.2%; Score 48; DB 1; Length 48;
Best Local Similarity 100.0%; Pred. No. 9.7e-05;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22582 GAGAGATAGCAACCGATCTTATATATCTACGTCACCCCTGCCCC 22629
|||
Db 48 GAGAGATAGCAACCGATCTTATATATCTACGTCACCCCTGCCCC 1

RESULT 2
ID T63871/c
AC T63871; 48 BP.

DT 14-MAR-1997 (first entry)
DE Human H4/a gene (for Histone H4) TFIIID binding site.
KW Duplex DNA; target region; binding characteristic; DNA binding protein;
KW TFIIID; transcription factor; binding site; inhibition; enhance;
KW cancer; inherited genetic disorder; alpha-D-galactosidase A; ds.
OS Homo sapiens.
PN US557844A.
PD 26-NOV-1996.
PF 27-JUN-1991; 723618.
PR 23-DEC-1992; US-966783.
PR 17-SEP-1993; US-123936.
PR 20-DEC-1993; US-171189.
PA (GENE-) GENELABS TECHNOLOGIES INC.
PI Andrews BM, Cantor CR, Edwards CA, Fry KE, Turin LM;
PI WPI: 97-020402/02.
PT Altering binding characteristics of DNA binding proteins to duplex
PT DNA - by attaching specific small cpd. to target region close to the
PT protein's binding site, useful in treatment of viral disease, cancer
PT etc.
PS Claim 6: Column 179-180; 264pp; English.
CC The sequences given in T63713-4312 represent duplex DNA's which act
CC as target regions in the method of the invention. The method for
CC altering the binding characteristics of a DNA-binding protein to duplex
CC DNA comprises contacting the duplex DNA with a small molecule which
CC binds sequence-specifically to a target region, where, when the small
CC molecule is bound to the target region, it is adjacent to, but not
CC overlapping by more than 4 bp, a binding site for a DNA-binding protein.
CC The small molecule is added at a concentration effective to alter the
CC binding of the DNA binding protein. The small molecule may inhibit or enhance
CC the binding of the DNA-binding protein to its binding site. The
CC compounds isolated using this method are potentially useful as
CC therapeutic agents for treatment of any disease which involves a
CC specific DNA sequence, e.g. cancer, or inherited genetic disorders etc.
CC The method is suitable for screening large biological or chemical
CC libraries and allows determination of sequence-specific and relative
CC affinities of known DNA-binding agents for different DNA sequences.
CC The design of these duplex DNA's allows a single DNA:protein interaction
CC to be used for screening sequence-specific, or preferential, DNA binding
CC proteins that recognise almost any possible sequence (see also T49539-
CC 74).
SQ Sequence 48 BP; 12 A; 8 C; 14 G; 14 T;
Query Match 0.2%; Score 48; DB 1; Length 48;
Best Local Similarity 100.0%; Pred. No. 9.7e-05;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 22582 GAGAGAAATGGAACCGATCTTATATCTACGTTACCCGCCGCC 22629
DB 48 GAGAGAAATGGAACCGATCTTATATCTACGTTACCCGCCGCC 1
RESULT 3
X17159/c
ID X17159 standard; DNA: 48 BP.
AC X17159;
DT 06-MAY-1999 (first entry)
DE Test sequence from human H4/a gene for H4 histone.
KW Test sequence; DNA-binding molecule; screening sequence; human;
KW nucleic acid amplification; target; viral; ds.
OS Homo sapiens.
PN US5869241-A.
PD 09-FEB-1999.
PF 07-JUN-1993; 475228.
PR 20-DEC-1993; US-171389.
PR 27-JUN-1991; US-723618.
PR 23-DEC-1992; US-966783.
PR 17-SEP-1993; US-123936.
PR 07-JUN-1993; US-475228.
PA (GENE-) GENELABS TECHNOLOGIES INC.
PI Andrews BM, Cantor CR, Edwards CA, Fry KE, Turin LM;

DR WPI: 99-152755/13.
PT Determination of DNA sequence preference of a DNA-binding molecule -
PT based on inhibition of binding of protein to oligonucleotide
PT sequence attached to test sequence
PS Claim 3: Columns 181-182; 270pp; English.
CC Sequences X17001 to X17600 represent specifically claimed target test
CC sequences that are used in the method of the invention of determining the
CC DNA sequence preference of a DNA-binding molecule. The method comprises:
CC (i) adding a test molecule and a DNA-binding protein to a mixture of
CC duplex DNA test oligonucleotides each of the test oligonucleotides
CC having a test sequence adjacent to a screening sequence, where the
CC screening sequence binds to the DNA-binding protein with a binding
CC affinity that is independent of the DNA sequence of the test sequence,
CC and where the mixture of duplex DNA test oligonucleotides includes
CC several test sequences; (ii) incubating the test molecule, the mixture of
CC duplex DNA test oligonucleotides and the DNA-binding protein for a time
CC sufficient to permit binding of the test molecule to test sequences in
CC the duplex DNA; (iii) separating unbound test oligonucleotides from test
CC oligonucleotides bound to binding protein; (iv) amplifying the unbound
CC test oligonucleotides; (v) repeating steps (ii) to (iv); (vi) isolating
CC the amplified test oligonucleotides; and (vii) sequencing the isolated
CC test oligonucleotides. Test sequences X17001-X17481 and X17600 correspond
CC to promoter targets for human genes and test sequences X17482-X17599
CC correspond to promoter targets for viral genes.
SQ Sequence 48 BP; 12 A; 8 C; 14 G; 14 T;
Query Match 0.2%; Score 48; DB 1; Length 48;
Best Local Similarity 100.0%; Pred. No. 9.7e-05;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 22582 GAGAGAAATGGAACCGATCTTATATCTACGTTACCCGCCGCC 22629
DB 48 GAGAGAAATGGAACCGATCTTATATCTACGTTACCCGCCGCC 1
RESULT 4
ID 069420 standard; DNA: 47 BP.
AC 069420;
DT 27-FEB-1995 (first entry)
DE Human H3 histone gene, target region.
KW DNA protein-binding assay; test sequence; screening sequence;
KW promoter; target; TATA box; Herpes Simplex Virus; HSV;
KW origin of replication; UL9; transcription factor; TFIIID; ds.
OS Synthetic.
PN MO9414980-A.
PD 07-JUL-1994.
PF 20-DEC-1993; U12388.
PR 23-DEC-1992; US-966783.
PR 17-SEP-1993; US-123936.
PA (GENE-) GENELABS TECHNOLOGIES INC.
PI Andrews BM, Cantor CR, Edwards CA, Fry KE, Turin LM;
PI WPI: 94-234711/28.
PT Sequence-directed DNA-binding molecules - useful in
PT pharmaceuticals and as molecular reagents
PS Claim 26: Page 29; 587pp; English.
CC A DNA protein-binding assay is provided, useful for screening
CC libraries of synthetic or biological cpds. for their ability
CC to bind DNA test sequences. The assay is versatile in that any
CC number of test sequences can be tested by placing the test sequence
CC adjacent to a defined protein-binding screening sequence. Binding
CC of moIs. to these test sequences changes the binding characteristics
CC of the protein mol. to its cognate binding sequence. When such a mol.
CC binds the test sequence, the equilibrium of the DNA:protein complexes
CC is disturbed, generating changes in the concentration of free DNA probe.
CC One application of this method is to eucaryotic general transcription
CC factors (e.g. TFIIID), where the target region is typically selected
CC from DNA sequences adjacent to the binding site for the eucaryotic
CC transcription factor. Numerous exemplary test sequences are given:
CC the sequences in 069251-731 and 069550 correspond to promoter targets
CC (typically TATA box-contg. sites) for human genes and the sequences in
CC 069732-849 correspond to promoter targets for viral genes. The test

CC sequences may also be randomly generated. DNA:protein interaction may
CC be used for screening purposes, e.g. the Herpes Simplex Virus (HSV)
CC origin of replication and US (see Q69851-52, Q69865 and Q69891).
SQ Sequence 47 BP; 13 A; 8 C; 10 G; 16 T;

Query Match 0.2%; Score 47; DB 1; Length 47;
Best Local Similarity 100.0%; Pred. No. 0.00018;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12078 CAATAGTGGTGGTCTGACTCTATAAAGAAGAGTACTCTTTCCCTT 12124
|||||
DB 1 CAATAGTGGTGGTCTGACTCTATAAAGAAGAGTACTCTTTCCCTT 47

RESULT 5
T63882 5
ID T63882 standard; DNA; 47 BP.
AC T63882;
DT 14-MAR-1997 (first entry)
DE Human H3 histone gene TFIID binding site.
KW Duplex DNA; target region; binding characteristic; DNA binding protein;
KW TFIID; transcription factor; binding site; inhibition; enhance; hcg;
KW cancer; inherited genetic disorder; alpha-D-galactosidase A; ds.
OS Homo sapiens.
PN US5578444-A.
PD 26-NOV-1996.
PF 27-JUN-1991; 723618.
PR 27-JUN-1991; US-723618.
PR 23-DEC-1992; US-996783.
PR 17-SEP-1993; US-123936.
PR 20-DEC-1993; US-171389.
PA (GENE-) GENELABS TECHNOLOGIES INC.
PI Andrews BM, Cantor CR, Edwards CA, Fry KE, Turin LM;
PT WPI: 97-020402/02.
PT Altering binding characteristics of DNA binding proteins to duplex
PT DNA - by attaching specific small cpd. to target region close to the
PT protein's binding site, useful in treatment of viral disease, cancer
PT etc

PS Claim 6; Column 185-186; 264pp; English.
CC The sequences given in T63713-4312 represent duplex DNA's which act
CC as target regions in the method of the invention. The method for
CC altering the binding characteristics of a DNA-binding protein to duplex
CC DNA comprises contacting the duplex DNA with a small molecule which
CC binds sequence-specifically to a target region, where, when the small
CC molecule is bound to the target region, it is adjacent to, but not
CC overlapping by more than 4 bp, a binding site for a DNA-binding protein.
CC The small molecule is added at a concentration effective to alter the
CC binding of the DNA binding protein, pref. TFIID, to its binding site on
CC the duplex DNA. The binding of the small molecule may inhibit or enhance
CC the binding of the DNA-binding protein to its binding site. The
CC compounds isolated using this method are potentially useful as
CC therapeutic agents for treatment of any disease which involves a
CC specific DNA sequence, e.g. cancer, or inherited genetic disorders etc.
CC The method is suitable for screening large biological or chemical
CC libraries and allows determination of sequence-specific and relative
CC affinities of known DNA-binding agents for different DNA sequences.
CC The design of these duplex DNA's allows a single DNA:protein interaction
CC to be used for screening sequence-specific, or preferential, DNA binding
CC proteins that recognise almost any possible sequence (see also T49539-
CC 74).
SQ Sequence 47 BP; 13 A; 8 C; 10 G; 16 T;

Query Match 0.2%; Score 47; DB 1; Length 47;
Best Local Similarity 100.0%; Pred. No. 0.00018;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12078 CAATAGTGGTGGTCTGACTCTATAAAGAAGAGTACTCTTTCCCTT 12124
|||||
DB 1 CAATAGTGGTGGTCTGACTCTATAAAGAAGAGTACTCTTTCCCTT 47

RESULT 6
X17170 6
ID X17170 standard; DNA; 47 BP.
AC X17170;
DT 06-MAY-1999 (first entry)
DE Test sequence from human histone H3 gene.
KW Test sequence: DNA-binding molecule; screening sequence; human;
KW nucleic acid amplification; target; viral; ds.
OS Homo sapiens.
PN US5869241-A.
PD 09-FEB-1999.
PF 07-JUN-1995; 475228.
PR 20-DEC-1993; US-171389.
PR 27-JUN-1991; US-723618.
PR 23-DEC-1992; US-996783.
PR 17-SEP-1993; US-123936.
PR 07-JUN-1995; US-475228.
PA (GENE-) GENELABS TECHNOLOGIES INC.
PI Andrews BM, Cantor CR, Edwards CA, Fry KE, Turin LM;
PT WPI: 99-152755/13.

PT Determination of DNA sequence preference of a DNA-binding molecule -
PT based on inhibition of binding of protein to oligonucleotide
PT sequence attached to test sequence
PS Claim 3; Columns 187-188; 270pp; English.
CC Sequences X17001 to X17600 represent specifically claimed target test
CC DNA sequence preference of a DNA-binding molecule. The method comprises:
CC (1) adding a test molecule and a DNA-binding protein to a mixture of
CC duplex DNA test oligonucleotides, each of the test oligonucleotides
CC having a test sequence adjacent to a screening sequence, where the
CC screening sequence binds to the DNA-binding protein with a binding
CC affinity that is independent of the DNA sequence of the test sequence,
CC and where the mixture of duplex DNA test oligonucleotides includes
CC several test sequences; (11) incubating the test molecule, the mixture of
CC duplex DNA test oligonucleotides and the DNA-binding protein for a time
CC sufficient to permit binding of the test molecule to test sequences in
CC the duplex DNA; (111) separating unbound test oligonucleotides from test
CC oligonucleotides bound to binding protein; (1v) amplifying the unbound
CC test oligonucleotides; (v) repeating steps (1) to (1v); (vi) isolating
CC the amplified test oligonucleotides; and (vii) sequencing the isolated
CC test oligonucleotides. Test sequences X17001-X17481 and X17600 correspond
CC to promoter targets for human genes and test sequences X17482-X17599
CC correspond to promoter targets for viral genes.
SQ Sequence 47 BP; 13 A; 8 C; 10 G; 16 T;

Query Match 0.2%; Score 47; DB 1; Length 47;
Best Local Similarity 100.0%; Pred. No. 0.00018;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12078 CAATAGTGGTGGTCTGACTCTATAAAGAAGAGTACTCTTTCCCTT 12124
|||||
DB 1 CAATAGTGGTGGTCTGACTCTATAAAGAAGAGTACTCTTTCCCTT 47

RESULT 7
Q69417 7
ID Q69417 standard; DNA; 46 BP.
AC Q69417;
DT 27-FEB-1995 (first entry)
DE Human H2A histone gene, target region.
KW DNA protein-binding assay; test sequence; screening sequence;
KW promoter; target; TATA box; Herpes Simplex Virus; HSV;
KW origin of replication; US; transcription factor; TFIID; ds.
OS Synthetic.
PN W09414980-A.
PD 07-JUL-1994.
PF 20-DEC-1993; U12388.
PR 23-DEC-1992; US-996783.
PR 17-SEP-1993; US-123936.
PA (GENE-) GENELABS TECHNOLOGIES INC.
PI Andrews BM, Cantor CR, Edwards CA, Fry KE, Turin LM;
PT WPI: 94-234711/28.

PT Sequence-directed DNA-binding molecules - useful in
PS pharmaceuticals and as molecular reagents
CC Claim 28, Page 295, 587pp; English.
CC A DNA protein-binding assay is provided, useful for screening
CC libraries of synthetic or biological cpts. for their ability
CC to bind DNA test sequences. The assay is versatile in that any
CC number of test sequences can be tested by placing the test sequence
CC adjacent to a defined protein-binding screening sequence. Binding
CC of mols. to these test sequences changes the binding characteristics
CC of the protein mol. to its cognate binding sequence. When such a mol.
CC binds the test sequence, the equilibrium of the DNA:protein complexes
CC is disturbed, generating changes in the concentration of free DNA probe.
CC One application of this method is to eucaryotic general transcription
CC factors (e.g., TFIID), where the target region is typically selected
CC from DNA sequences adjacent to the binding site for the eucaryotic
CC transcription factor. Numerous exemplary test sequences are given:
CC the sequences in 069251-731 and 069850 correspond to promoter targets
CC (typically, TATA box-contg. sites) for human genes and the sequences in
CC 069732-849 correspond to promoter targets for viral genes. The test
CC sequences may also be randomly generated. DNA:protein interaction may
CC be used for screening purposes, e.g., the Herpes Simplex Virus (HSV)
CC origin of replication and US9 (see 069851-52, 069855 and 069891).
SQ Sequence 46 BP; 10 A; 16 C; 6 G; 14 T;

Query Match 0.2%; Score 46; DB 1; Length 46;
Best Local Similarity 100.0%; Pred. No. 0.00034;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10562 CACAGCCTACCTCCAGTCAGTATATAACTCTCTGCTTCGCTTC 10607
DB 1 CACAGCCTACCTCCAGTCAGTATATAACTCTCTGCTTCGCTTC 46

RESULT 8
TS3879 163879 standard; DNA; 46 BP.
AC 163879:
DE Human H2A histone gene TFIID binding site.
KW Duplex DNA; target region; binding characteristic; DNA binding protein;
KW TFIID; transcription factor; binding site; inhibition; enhance; hcg;
KW cancer; inherited genetic disorder; alpha-D-galactosidase A; ds.
OS Homo sapiens.
PN US5578444-A.
PD 26-NOV-1996.
PR 27-JUN-1991; US-723618.
PR 27-JUN-1991; US-723618.
PR 23-DEC-1992; US-996783.
PR 17-SEP-1993; US-123936.
PR 20-DEC-1993; US-171389.
PI (GENE-) GENELABS TECHNOLOGIES INC.
PI Andrews BM, Cantor CR, Edwards CA, Fry KE, Turin LM;
DR WPI: 99-020402/02.
PT Altering binding characteristics of DNA binding proteins to duplex
PT DNA - by attaching specific small cpd. to target region close to the
PT protein's binding site, useful in treatment of viral disease, cancer
PT etc.
PS Claim 6; Column 183-184; 264pp; English.
CC The sequences given in T63713-4312 represent duplex DNA's which act
CC as target regions in the method of the invention. The method for
CC altering the binding characteristics of a DNA-binding protein to duplex
CC DNA comprises contacting the duplex DNA with a small molecule which
CC binds sequence-specifically to a target region, where, when the small
CC molecule is bound to the target region, it is adjacent to, but not
CC overlapping by more than 4 bp, a binding site for a DNA-binding protein.
CC The small molecule is added at a concentration effective to alter the
CC binding of the DNA binding protein, pref. TFIID, to its binding site on
CC the duplex DNA. The binding of the small molecule may inhibit or enhance
CC the binding of the DNA-binding protein to its binding site. The
CC compounds isolated using this method are potentially useful as
CC therapeutic agents for treatment of any disease which involves a
CC specific DNA sequence, e.g. cancer, or inherited genetic disorders etc.

CC The method is suitable for screening large biological or chemical
CC libraries and allows determination of sequence-specific and relative
CC affinities of known DNA-binding agents for different DNA sequences.
CC The design of these duplex DNA's allows a single DNA:protein interaction
CC to be used for screening sequence-specific, or preferential, DNA binding
CC proteins that recognise almost any possible sequence (see also T63539-
CC 74).
SQ Sequence 46 BP; 10 A; 16 C; 6 G; 14 T;

Query Match 0.2%; Score 46; DB 1; Length 46;
Best Local Similarity 100.0%; Pred. No. 0.00034;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10562 CACAGCCTACCTCCAGTCAGTATATAACTCTCTGCTTCGCTTC 10607
DB 1 CACAGCCTACCTCCAGTCAGTATATAACTCTCTGCTTCGCTTC 46

RESULT 9
X17167 163167 standard; DNA; 46 BP.
AC X17167:
DE 06-MAY-1999 (first entry)
KW Test sequence from human histone H2a gene.
KW nucleic acid amplification; target; viral; ds.
OS Homo sapiens.
PN US5869241-A.
PD 09-FEB-1999.
PR 07-JUN-1993; US-475228.
PR 20-DEC-1993; US-171389.
PR 27-JUN-1991; US-723618.
PR 23-DEC-1992; US-996783.
PR 17-SEP-1993; US-123936.
PR 07-JUN-1993; US-475228.
PI (GENE-) GENELABS TECHNOLOGIES INC.
PI Andrews BM, Cantor CR, Edwards CA, Fry KE, Turin LM;
DR WPI: 99-152755/13.
PT Determination of DNA sequence preference of a DNA-binding molecule -
PT based on inhibition of binding of protein to oligonucleotide
PT sequence attached to test sequence
PS Claim 3; Columns 185-186; 270pp; English.
CC Sequences X17001 to X17600 represent specifically claimed target test
CC DNA sequences that are used in the method of the invention of determining the
CC DNA sequence preference of a DNA-binding molecule. The method comprises:
CC (1) adding a test molecule and a DNA-binding protein to a mixture of
CC duplex DNA test oligonucleotides, each of the test oligonucleotides
CC having a test sequence adjacent to a screening sequence, where the
CC screening sequence binds to the DNA-binding protein with a binding
CC affinity that is independent of the DNA sequence of the test sequence,
CC and where the mixture of duplex DNA test oligonucleotides includes
CC several test sequences; (11) incubating the test molecule, the mixture of
CC duplex DNA test oligonucleotides and the DNA-binding protein for a time
CC sufficient to permit binding of the test molecule to test sequences in
CC the duplex DNA; (111) separating unbound test oligonucleotides from test
CC oligonucleotides bound to binding protein; (1V) amplifying the unbound
CC test oligonucleotides; (V) repeating steps (11) to (1V); (VI) isolating
CC the amplified test oligonucleotides; and (VII) sequencing the isolated
CC test oligonucleotides. Test sequences X17001-X17481 and X17600 correspond
CC to promoter targets for human genes and test sequences X17482-X17599
CC correspond to promoter targets for viral genes.
SQ Sequence 46 BP; 10 A; 16 C; 6 G; 14 T;

Query Match 0.2%; Score 46; DB 1; Length 46;
Best Local Similarity 100.0%; Pred. No. 0.00034;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10562 CACAGCCTACCTCCAGTCAGTATATAACTCTCTGCTTCGCTTC 10607
DB 1 CACAGCCTACCTCCAGTCAGTATATAACTCTCTGCTTCGCTTC 46

```
RESULT 10
ID 069419 standard; DNA: 45 BP.
AC 069419;
DE 27-FEB-1995 (first entry)
KW Human H3 histone gene, target region.
KW DNA protein-binding assay; test sequence; screening sequence;
KW promoter; target; TATA box; Herpes Simplex Virus; HSV;
KW origin of replication; UL9; transcription factor; TFIID; ds.
OS Synthetic.
PN WO9414980-A.
PD 07-JUL-1994.
PF 20-DEC-1993; U12388.
PR 23-DEC-1992; US-996783.
PR 17-SEP-1993; US-123936.
PA (GENE-) GENELABS TECHNOLOGIES INC.
PI Andrews BM, Cantor CR, Edwards CA, Fry KE, Turin LM;
DR WPI: 94-234711/28
PT Sequence-directed DNA-binding molecules - useful in
PT pharmaceuticals and as molecular reagents
PS Claim 28; Page 296; 587pp; English.
CC A DNA protein-binding assay is provided, useful for screening
CC libraries of synthetic or biological cpds. for their ability
CC to bind DNA test sequences. The assay is versatile in that any
CC number of test sequences can be tested by placing the test sequence
CC adjacent to a defined protein-binding screening sequence. Binding
CC of mols. to these test sequences changes the binding characteristics
CC of the protein mol. to its cognate binding sequence. When such a mol.
CC binds the test sequence, the equilibrium of the DNA:protein complexes
CC is disturbed, generating changes in the concentration of free DNA probe.
CC One application of this method is to eucaryotic general transcription
CC factors (e.g. TFIID) where the target region is typically selected
CC from DNA sequences adjacent to the binding site for the eucaryotic
CC transcription factor. Numerous exemplary test sequences are given:
CC the sequences in 069251-731 and 069850 correspond to promoter targets
CC (typically, TATA box-contg. sites) for human genes and the sequences in
CC 069732-849 correspond to promoter targets for viral genes. The test
CC sequences may also be randomly generated. DNA:protein interaction may
CC be used for screening purposes, e.g. the Herpes Simplex Virus (HSV)
CC origin of replication and UL9 (see 069851-52; 069865 and 069891).
SQ Sequence 45 BP; 8 A; 6 C; 10 G; 21 T;

Query Match 0.1%; Score 45; DB 1; Length 45;
Best Local Similarity 100.0%; Pred. No. 0.00065;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12057 ATTTTGAATTTCTTGCGTCCAAATAGTGTGCTGACTCTAT 12101
DB 1 ATTTTGAATTTCTTGCGTCCAAATAGTGTGCTGACTCTAT 45

RESULT 11
ID T63881 standard; DNA: 45 BP.
AC T63881;
DE 14-MAR-1997 (first entry)
KW Human H3 histone gene TFIID binding site.
KW Duplex DNA; target region; binding characteristic; DNA binding protein;
KW TFIID; transcription factor; binding site; inhibition; enhance; hCG;
KW cancer; inherited genetic disorder; alpha-D-galactosidase A; ds.
OS Homo sapiens.
PN US5578444-A.
PD 26-NOV-1996.
PF 27-JUN-1991; 723618.
PR 27-JUN-1991; US-723618.
PR 23-DEC-1992; US-996783.
PR 17-SEP-1993; US-123936.
PR 20-DEC-1993; US-171389.
PA (GENE-) GENELABS TECHNOLOGIES INC.
PI Andrews BM, Cantor CR, Edwards CA, Fry KE, Turin LM;
DR WPI: 97-020402/02.
```

```
PT Altering binding characteristics of DNA binding proteins to duplex
PT DNA - by attaching specific small cpd. to target region close to the
PT protein's binding site, useful in treatment of viral disease, cancer
PT etc
PS Claim 6; Column 185-186; 264pp; English.
CC The sequences given in T63713-4312 represent duplex DNA's which act
CC as target regions in the method of the invention. The method for
CC altering the binding characteristics of a DNA-binding protein to duplex
CC DNA comprises contacting the duplex DNA with a small molecule which
CC binds sequence-specifically to a target region, where, when the small
CC molecule is bound to the target region, it is adjacent to, but not
CC overlapping by more than 4 bp, a binding site for a DNA-binding protein.
CC The small molecule is added at a concentration effective to alter the
CC binding of the DNA binding protein, pref. TFIID, to its binding site on
CC the duplex DNA. The binding of the small molecule may inhibit or enhance
CC the binding of the DNA-binding protein to its binding site. The
CC compounds isolated using this method are potentially useful as
CC therapeutic agents for treatment of any disease which involves a
CC specific DNA sequence, e.g. cancer, or inherited genetic disorders etc.
CC The method is suitable for screening large biological or chemical
CC libraries and allows determination of sequence-specific and relative
CC affinities of known DNA-binding agents for different DNA sequences.
CC The design of these duplex DNA's allows a single DNA:protein interaction
CC to be used for screening sequence-specific, or preferential, DNA binding
CC proteins that recognise almost any possible sequence (see also T49539-
CC 74).
SQ Sequence 45 BP; 8 A; 6 C; 10 G; 21 T;

Query Match 0.1%; Score 45; DB 1; Length 45;
Best Local Similarity 100.0%; Pred. No. 0.00065;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12057 ATTTTGAATTTCTTGCGTCCAAATAGTGTGCTGACTCTAT 12101
DB 1 ATTTTGAATTTCTTGCGTCCAAATAGTGTGCTGACTCTAT 45

RESULT 12
ID X17169 standard; DNA: 45 BP.
AC X17169;
DE 06-MAY-1999 (first entry)
KW Test sequence; DNA-binding molecule; screening sequence; human;
KW nucleic acid amplification; target; viral; ds.
OS Homo sapiens.
PN US5869241-A.
PD 09-FEB-1999.
PF 07-JUN-1995; 475228.
PR 20-DEC-1993; US-171389.
PR 27-JUN-1991; US-723618.
PR 23-DEC-1992; US-996783.
PR 17-SEP-1993; US-123936.
PR 07-JUN-1995; US-475228.
PA (GENE-) GENELABS TECHNOLOGIES INC.
PI Andrews BM, Cantor CR, Edwards CA, Fry KE, Turin LM;
DR WPI: 99-152735/13.
PT Determination of DNA sequence preference of a DNA-binding molecule -
PT based on inhibition of binding of protein to oligonucleotide
PT sequence attached to test sequence
PS Claim 3; Columns 185-186; 270pp; English.
CC Sequences X17001 to X17600 represent specifically claimed target test
CC sequences that are used in the method of the invention of determining the
CC DNA sequence preference of a DNA-binding molecule. The method comprises:
CC (1) adding a test molecule and a DNA-binding protein to a mixture of
CC duplex DNA test oligonucleotides, each of the test oligonucleotides
CC having a test sequence adjacent to a screening sequence, where the
CC screening sequence binds to the DNA-binding protein with a binding
CC affinity that is independent of the DNA sequence of the test sequence,
CC and where the mixture of duplex DNA test oligonucleotides includes
CC several test sequences; (11) incubating the test molecule, the mixture of
CC duplex DNA test oligonucleotides and the DNA-binding protein for a time
```

CC sufficient to permit blinding of the test molecule to test sequences in
CC the duplex DNA; (iii) separating unbound test oligonucleotides from test
CC oligonucleotides bound to binding protein; (iv) amplifying the unbound
CC test oligonucleotides; (v) repeating steps (ii) to (iv); (vi) isolating
CC the amplified test oligonucleotides; and (vii) sequencing the isolated
CC test oligonucleotides. Test sequences X17001-X17481 and X17600 correspond
CC to promoter targets for human genes and test sequences X17482-X17599
CC correspond to promoter targets for viral genes
SQ Sequence 45 BP; 8 A; 6 C; 10 G; 21 T;

Query Match 0.1%; Score 45; DB 1; Length 45;
Best Local Similarity 100.0%; Pred. No. 0.00065;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12057 ATTTTGAATTTCTTGGTCCAAATGTTGTTGCTGACTCTAT 12101
DB 1 ATTTTGAATTTCTTGGTCCAAATGTTGTTGCTGACTCTAT 45

RESULT 13
062216/c
ID 062216 standard; DNA; 30 BP.
AC 062216;
DT 22-NOV-1994 (first entry)
DE Histone H3 mRNA probe #1
KW Probe; histone; H2B; H3; H4; mRNA; detection; proliferation; cell;
OS Synthetic;
PN WO9408046-A.
PD 14-APR-1994;
PF 20-SEP-1993; GB-020777.
PR 02-OCT-1992; GB-020777.
PS (UTILE-) UNIV LEICESTER.
PT Pringle JH;
PI WPI; 94-135606/16.
DR Detecting proliferating cells in tissue or cell samples - by
PT hybridisation with non-radioactively labelled oligonucleotide
PT probes specific for histone mRNA.
PS Claim 13; Page 18; 24pp; English.
CC The sequences given in 062216-18 are probes which are specific for
CC histone H3 mRNA. These probes may be used in the method of the
CC invention for detecting the presence of proliferating cells in a
CC tissue/cell sample. The method comprises preparing the tissue/cell
CC sample, hybridising these non-radioactively labelled probes, and
CC detecting the presence of histone-probe hybrids. The method can be
CC used to identify cells undergoing proliferation in normal and
CC diseased tissue. It can be used in the diagnosis of cancers and in
CC assessing the carcinogenic properties of a chemical.
SQ Sequence 30 BP; 6 A; 6 C; 8 G; 10 T;

Query Match 0.1%; Score 30; DB 1; Length 30;
Best Local Similarity 100.0%; Pred. No. 9;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12518 CGATGACTATTTATGCCCAAGACATCCAG 12547
DB 30 CGATGACTATTTATGCCCAAGACATCCAG 1

RESULT 14
062217/c
ID 062217 standard; DNA; 30 BP.
AC 062217;
DT 22-NOV-1994 (first entry)
DE Histone H3 mRNA probe #2.
KW Probe; histone; H2B; H3; H4; mRNA; detection; proliferation; cell;
OS Synthetic;
PN WO9408046-A.
PD 14-APR-1994;
PF 20-SEP-1993; GB-01977.
Job time: 136132 sec

PR 02-OCT-1992; GB-020777.
PA (UTILE-) UNIV LEICESTER.
PI Pringle JH;
DR WPI; 94-135606/16.
PT Detecting proliferating cells in tissue or cell samples - by
PT hybridisation with non-radioactively labelled oligonucleotide
PT probes specific for histone mRNA.
PS Claim 13; Page 18; 24pp; English.
CC The sequences given in 062216-18 are probes which are specific for
CC histone H3 mRNA. These probes may be used in the method of the
CC invention for detecting the presence of proliferating cells in a
CC tissue/cell sample. The method comprises preparing the tissue/cell
CC sample, hybridising these non-radioactively labelled probes, and
CC detecting the presence of histone-probe hybrids. The method can be
CC used to identify cells undergoing proliferation in normal and
CC diseased tissue. It can be used in the diagnosis of cancers and in
CC assessing the carcinogenic properties of a chemical.
SQ Sequence 30 BP; 7 A; 7 C; 8 G; 8 T;

Query Match 0.1%; Score 30; DB 1; Length 30;
Best Local Similarity 100.0%; Pred. No. 9;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12330 GCTACCAAAAGTCGACGAGTTCTGATTC 12359
DB 30 GCTACCAAAAGTCGACGAGTTCTGATTC 1

RESULT 15
062218/c
ID 062218 standard; DNA; 30 BP.
AC 062218;
DT 22-NOV-1994 (first entry)
DE Histone H3 mRNA probe #3.
KW Probe; histone; H2B; H3; H4; mRNA; detection; proliferation; cell;
OS Synthetic;
PN WO9408046-A.
PD 14-APR-1994;
PF 20-SEP-1993; GB-020777.
PR 02-OCT-1992; GB-020777.
PS (UTILE-) UNIV LEICESTER.
PT Pringle JH;
PI WPI; 94-135606/16.
DR Detecting proliferating cells in tissue or cell samples - by
PT hybridisation with non-radioactively labelled oligonucleotide
PT probes specific for histone mRNA.
PS Claim 13; Page 18; 24pp; English.
CC The sequences given in 062216-18 are probes which are specific for
CC histone H3 mRNA. These probes may be used in the method of the
CC invention for detecting the presence of proliferating cells in a
CC tissue/cell sample. The method comprises preparing the tissue/cell
CC sample, hybridising these non-radioactively labelled probes, and
CC detecting the presence of histone-probe hybrids. The method can be
CC used to identify cells undergoing proliferation in normal and
CC diseased tissue. It can be used in the diagnosis of cancers and in
CC assessing the carcinogenic properties of a chemical.
SQ Sequence 30 BP; 4 A; 6 C; 9 G; 11 T;

Query Match 0.1%; Score 30; DB 1; Length 30;
Best Local Similarity 100.0%; Pred. No. 9;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12177 GTACTAAACAGACAGCTCGGAATCCACCG 12206
DB 30 GTACTAAACAGACAGCTCGGAATCCACCG 1

Search completed: March 30, 2000, 17:38:54
Job time: 136132 sec

Mon Apr 3 08:24:32 2000

us-08-852-495c-1_copy_90000_120000.png

Page 7



▼
:
6

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 30, 2000, 10:42:50 ; Search time 344.26 Seconds
(without alignments)
10433.435 Million cell updates/sec

Title: US-08-852-495C-1_COPY_90000_120000

Perfect score: 30001
Sequence: 1 TGGCGAGCTGAGGTGGCGG.....TATCTAGCATGTCAT 30001

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 214294 seqs, 59861574 residues

Word size : 0

Total number of hits satisfying chosen parameters: 349160

Minimum DB seq length: 8
Maximum DB seq length: 105

Post-processing: Listing first 45 summaries

Database :

1: /cgn2_6/prodata/1/lna/5A_COMB.seq:*
2: /cgn2_6/prodata/1/lna/5B_COMB.seq:*
3: /cgn2_6/prodata/1/lna/5C_COMB.seq:*
4: /cgn2_6/prodata/1/lna/5D_COMB.seq:*
5: /cgn2_6/prodata/1/lna/6_COMB.seq:*
6: /cgn2_6/prodata/1/lna/PCTUS9_COMB.seq:*
7: /cgn2_6/prodata/1/lna/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	58	0.2	76	3	US-08-454-557C-69
2	58	0.2	76	4	US-08-340-426D-69
3	58	0.2	76	4	US-08-450-673C-69
4	58	0.2	76	6	PCT-US95-17111A-69
5	55	0.2	60	3	US-08-454-557C-57
6	55	0.2	60	4	US-08-340-426D-57
7	55	0.2	60	4	US-08-450-673C-57
8	55	0.2	60	6	PCT-US95-17111A-57
9	48	0.2	48	1	US-08-171-389-159
10	48	0.2	48	2	US-08-123-936-159
11	48	0.2	48	3	US-08-475-228A-159
12	48	0.2	48	5	US-08-482-080A-159
13	48	0.2	48	6	PCT-US93-12388-159
14	47	0.2	47	1	US-08-171-389-170
15	47	0.2	47	2	US-08-123-936-170
16	47	0.2	47	3	US-08-475-228A-170
17	47	0.2	47	5	US-08-482-080A-170
18	47	0.2	47	6	PCT-US93-12388-170
19	46	0.2	46	1	US-08-171-389-167
20	46	0.2	46	2	US-08-123-936-167
21	46	0.2	46	3	US-08-475-228A-167
22	46	0.2	46	5	US-08-482-080A-167
23	46	0.2	46	6	PCT-US93-12388-167
24	45	0.1	45	1	US-08-171-389-169
25	45	0.1	45	2	US-08-123-936-169
26	45	0.1	45	3	US-08-475-228A-169

27	45	0.1	45	5	US-08-482-080A-169	Sequence 169, App
28	45	0.1	45	6	PCT-US93-12388-169	Sequence 169, App
29	28	0.1	30	4	US-08-859-998-318	Sequence 318, App
30	28	0.1	66	3	US-08-454-557C-67	Sequence 67, App1
31	28	0.1	66	4	US-08-340-426D-67	Sequence 67, App1
32	28	0.1	66	4	US-08-450-673C-67	Sequence 67, App1
33	28	0.1	66	6	PCT-US95-17111A-67	Sequence 67, App1
34	28	0.1	85	3	US-08-454-557C-92	Sequence 92, App1
35	28	0.1	85	4	US-08-340-426D-92	Sequence 92, App1
36	28	0.1	85	4	US-08-450-673C-92	Sequence 92, App1
37	28	0.1	85	6	PCT-US95-17111A-92	Sequence 92, App1
38	27	0.1	36	3	US-08-669-721-5	Sequence 5, App1
39	27	0.1	57	1	US-08-222-177A-163	Sequence 163, App1
40	27	0.1	84	3	US-08-454-557C-91	Sequence 91, App1
41	27	0.1	84	4	US-08-340-426D-91	Sequence 91, App1
42	27	0.1	84	4	US-08-450-673C-91	Sequence 91, App1
43	27	0.1	84	6	PCT-US95-17111A-91	Sequence 91, App1
44	26	0.1	35	1	US-08-255-889-10	Sequence 10, App1
45	26	0.1	36	3	US-08-669-721-5	Sequence 5, App1

ALIGNMENTS

RESULT 1
US-08-454-557C-69
Sequence 69, Application US/08454557C
Patent No. 5830670
GENERAL INFORMATION:
APPLICANT: de la Monte, Suzanne
TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection
TITLE OF INVENTION: of Alzheimer's Disease
NUMBER OF SEQUENCES: 121
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/454,557C
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609,3840003
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 69:
SEQUENCE CHARACTERISTICS:
LENGTH: 76 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
US-08-454-557C-69
Query Match 0.2%: Score 58; DB 3; Length 76;
Best Local Similarity 100.0%: Pred. No. 2.1e+10;
Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 2685 GCCCGCTAATTTTATTTAGTAGACACAGGCTTCACCGTGTGGCCAGATG 2742
DB 6 GCCCGCTAATTTTATTTAGTAGACACAGGCTTCACCGTGTGGCCAGATG 63

Mon Apr 3 08:24:33 2000

us-08-852-495c-1_copy_90000_120000.rml

Page 2

```
RESULT 2
US-08-340-426D-69
Sequence 69, Application US/08340426D
Patent No. 5948634
GENERAL INFORMATION:
APPLICANT: de la Monte, Suzanne
TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection
NUMBER OF SEQUENCES: 121
CORRESPONDENCE ADDRESS:
ADDRESS: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/340,426D
FILING DATE: 14-NOV-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609.3840002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2500
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 69:
SEQUENCE CHARACTERISTICS:
LENGTH: 76 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
US-08-340-426D-69

Query Match          0.2%: Score 58; DB 4; Length 76;
Best Local Similarity 100.0%; Pred. No. 2.1e-10;
Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 2685 GCCCGCGTAATTTTGTATTTAGTAGAGACAGGCTTCCACCGTGGCCAGAGG 2742
DB 6 GCCCGCGTAATTTTGTATTTAGTAGAGACAGGCTTCCACCGTGGCCAGAGG 63

RESULT 3
US-08-450-673C-69
Sequence 69, Application US/08450673C
Patent No. 5948888
GENERAL INFORMATION:
APPLICANT: de la Monte, Suzanne
TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection
NUMBER OF SEQUENCES: 121
CORRESPONDENCE ADDRESS:
ADDRESS: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
```

```
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/450,673C
FILING DATE: 30-MAY-1995
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609.3840004
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2500
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 69:
SEQUENCE CHARACTERISTICS:
LENGTH: 76 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
US-08-450-673C-69

Query Match          0.2%: Score 58; DB 4; Length 76;
Best Local Similarity 100.0%; Pred. No. 2.1e-10;
Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 2685 GCCCGCGTAATTTTGTATTTAGTAGAGACAGGCTTCCACCGTGGCCAGAGG 2742
DB 6 GCCCGCGTAATTTTGTATTTAGTAGAGACAGGCTTCCACCGTGGCCAGAGG 63

RESULT 4
PCT-US95-17111A-69
Sequence 69, Application PC/TUS9517111A
GENERAL INFORMATION:
APPLICANT: de la Monte, Suzanne
TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection
NUMBER OF SEQUENCES: 121
CORRESPONDENCE ADDRESS:
ADDRESS: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/17111A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/340,426
FILING DATE: 14-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609.3840002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2540
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 69:
SEQUENCE CHARACTERISTICS:
LENGTH: 76 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
PCT-US95-17111A-69
```

Query Match 0.2%; Score 58; DB 6; Length 76;
Best Local Similarity 100.0%; Pred. No. 2.1e-10;
Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2685 GCCCCGCTAATTTTGTATTTAGTAGAGACAGGGTTCCACCGTGTGGCCAGATG 2742
DB 6 GCCCCGCTAATTTTGTATTTAGTAGAGACAGGGTTCCACCGTGTGGCCAGATG 63

RESULT 5

US-08-454-557C-57
; Sequence 57, Application US/08454557C
; Patent No. 5830670
; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; TITLE OF INVENTION: Neutral Thread Protein Gene Expression and Detection
; TITLE OF INVENTION: of Alzheimer's Disease
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/454,557C
; FILING DATE: 30-MAY-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Ludwig, Steven R.
; REGISTRATION NUMBER: 36,203
; REFERENCE/DOCKET NUMBER: 0609.3840003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 57:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 60 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
US-08-454-557C-57

Query Match 0.2%; Score 55; DB 3; Length 60;
Best Local Similarity 100.0%; Pred. No. 2e-09;
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2685 GCCCCGCTAATTTTGTATTTAGTAGAGACAGGGTTCCACCGTGTGGCCAGG 2739
DB 6 GCCCCGCTAATTTTGTATTTAGTAGAGACAGGGTTCCACCGTGTGGCCAGG 60

RESULT 6

US-08-340-426D-57
; Sequence 57, Application US/08340426D
; Patent No. 5948634
; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; TITLE OF INVENTION: Neutral Thread Protein Gene Expression and Detection
; TITLE OF INVENTION: of Alzheimer's Disease
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESSES:

ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/340,426D
; FILING DATE: 14-NOV-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Ludwig, Steven R.
; REGISTRATION NUMBER: 36,203
; REFERENCE/DOCKET NUMBER: 0609.3840002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 57:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 60 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
US-08-340-426D-57

Query Match 0.2%; Score 55; DB 4; Length 60;
Best Local Similarity 100.0%; Pred. No. 2e-09;
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2685 GCCCCGCTAATTTTGTATTTAGTAGAGACAGGGTTCCACCGTGTGGCCAGG 2739
DB 6 GCCCCGCTAATTTTGTATTTAGTAGAGACAGGGTTCCACCGTGTGGCCAGG 60

RESULT 7

US-08-450-673C-57
; Sequence 57, Application US/08450673C
; Patent No. 5948888
; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; TITLE OF INVENTION: Neutral Thread Protein Gene Expression and Detection
; TITLE OF INVENTION: of Alzheimer's Disease
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/450,673C
; FILING DATE: 30-MAY-1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Ludwig, Steven R.
; REGISTRATION NUMBER: 36,203
; REFERENCE/DOCKET NUMBER: 0609.3840004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540

Mon Apr 3 08:24:33 2000

us-08-852-495c-1_copy_90000_120000.rn1

Page 4

;; INFORMATION FOR SEQ ID NO: 57;
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 60 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: both
;; TOPOLOGY: both
US-08-450-673C-57

Query Match 0.2%; Score 55; DB 4; Length 60;
Best Local Similarity 100.0%; Pred. No. 2e-09;
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2685 GCCCGCGTAATTTTGTATTTAGTAGAGACAGGGTTTCACCGTGTGGCCAGG 2739
DB 6 GCCCGCGTAATTTTGTATTTAGTAGAGACAGGGTTTCACCGTGTGGCCAGG 60

RESULT 8
PCT-US95-17111A-57
Sequence 57, Application PC/TUS9517111A
GENERAL INFORMATION:
APPLICANT: de la Monte, Susanne
TITLE OF INVENTION: Neutral Thread Protein Gene Expression and
NUMBER OF SEQUENCES: 121
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/17111A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/340,426
FILING DATE: 14-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lusk, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609.3840002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2540
TELEFAX: (202) 371-2600
INFORMATION FOR SEQ ID NO: 57:
SEQUENCE CHARACTERISTICS:
LENGTH: 60 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
PCT-US95-17111A-57

Query Match 0.2%; Score 55; DB 6; Length 60;
Best Local Similarity 100.0%; Pred. No. 2e-09;

Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2685 GCCCGCGTAATTTTGTATTTAGTAGAGACAGGGTTTCACCGTGTGGCCAGG 2739
DB 6 GCCCGCGTAATTTTGTATTTAGTAGAGACAGGGTTTCACCGTGTGGCCAGG 60

RESULT 9
US-08-171-389-159/C

;; Sequence 159, Application US/08171389
;; Patent No. 5578444
;; GENERAL INFORMATION:
;; APPLICANT: Edwards, Cynthia A.
;; APPLICANT: Cantor, Charles R.
;; APPLICANT: Andrews, Beth M.
;; APPLICANT: Turin, Lisa M.
;; APPLICANT: Fry, Kirk E.
;; TITLE OF INVENTION: Sequence-directed DNA Binding
;; NUMBER OF SEQUENCES: 641
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Genelabs Technologies, Inc.
;; STREET: 505 Penobscot Drive
;; CITY: Redwood City
;; STATE: CA
;; COUNTRY: USA
;; ZIP: 94063
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/171,389
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/123,936
FILING DATE: 17-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/996,783
FILING DATE: 23-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/723,618
FILING DATE: 27-JUN-1991
APPLICATION NUMBER: US 08/081,070
FILING DATE: 22-JUN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Fadian, Gary R.
REGISTRATION NUMBER: 33,875
REFERENCE/DOCKET NUMBER: 4600-0175/G19P3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 159:
SEQUENCE CHARACTERISTICS:
LENGTH: 48 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: Human H4/a gene for H4 histone
US-08-171-389-159

Query Match 0.2%; Score 48; DB 1; Length 48;
Best Local Similarity 100.0%; Pred. No. 4.1e-07;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 22582 GAGAGATAGGAAGACGATCTTATATCTAGTACCGCCGCCCC 22629
DB 48 GAGAGATAGGAAGACGATCTTATATCTAGTACCGCCGCCCC 1

RESULT 10
US-08-123-936-159/C
Sequence 159, Application US/08123936
Patent No. 5726014
GENERAL INFORMATION:

APPLICANT: Edwards, Cynthia A.
APPLICANT: Cantor, Charles R.
APPLICANT: Andrews, Beth M.
APPLICANT: Turin, Lisa M.
TITLE OF INVENTION: Screening Assay for the Detection of
NUMBER OF SEQUENCES: 640
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genelabs Technologies, Inc.
STREET: 505 Penobscot Drive
CITY: Redwood City
STATE: CA
COUNTRY: USA
ZIP: 94063
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/123,936
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/996,783
FILING DATE: 23-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/723,618
FILING DATE: 27-JUN-1991
ATTORNEY/AGENT INFORMATION:
NAME: Fabian, Gary R.
REGISTRATION NUMBER: 33,875
REFERENCE/DOCKET NUMBER: 4600-0075.32/G19P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ. ID NO: 159:
SEQUENCE CHARACTERISTICS:
LENGTH: 48 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: Human H4/a gene for H4 histone
US-08-123-936-159

Query Match 0.2%; Score 48; DB 2; Length 48;
Best Local Similarity 100.0%; Pred. No. 4.1e-07;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22582 GAGAGATAGGAACGATCTTATATCTACGTTACCCCTGCCCC 22629
|||||
DB 48 GAGAGATAGGAACGATCTTATATCTACGTTACCCCTGCCCC 1

RESULT 11
US-08-475-228A-159/c
Sequence 159, Application US/08475228A
Patent No. 5869241
GENERAL INFORMATION:
APPLICANT: Edwards, Cynthia A.
APPLICANT: Cantor, Charles R.
APPLICANT: Andrews, Beth M.
APPLICANT: Turin, Lisa M.
APPLICANT: Fiy, Kirk E.
TITLE OF INVENTION: Sequence-Directed DNA Binding
TITLE OF INVENTION: Molecules, Compositions and Methods
NUMBER OF SEQUENCES: 664
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genelabs Technologies, Inc.

STREET: 505 Penobscot Drive
CITY: Redwood City
STATE: CA
COUNTRY: USA
ZIP: 94063
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/475,228A
FILING DATE: 06-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/123,936
FILING DATE: 17-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/996,783
FILING DATE: 23-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/723,618
FILING DATE: 27-JUN-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/081,070
FILING DATE: 22-JUN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Stratford, Carol A.
REGISTRATION NUMBER: 34,444
REFERENCE/DOCKET NUMBER: 4600-0175.21/G19P3D2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ. ID NO: 159:
SEQUENCE CHARACTERISTICS:
LENGTH: 48 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: Human H4/a gene for H4 histone
US-08-475-228A-159

Query Match 0.2%; Score 48; DB 3; Length 48;
Best Local Similarity 100.0%; Pred. No. 4.1e-07;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22582 GAGAGATAGGAACGATCTTATATCTACGTTACCCCTGCCCC 22629
|||||
DB 48 GAGAGATAGGAACGATCTTATATCTACGTTACCCCTGCCCC 1

RESULT 12
US-08-482-080A-159/c
Sequence 159, Application US/08482080A
Patent No. 6010849
GENERAL INFORMATION:
APPLICANT: Edwards, Cynthia A.
APPLICANT: Cantor, Charles R.
APPLICANT: Andrews, Beth M.
APPLICANT: Turin, Lisa M.
APPLICANT: Fiy, Kirk E.
TITLE OF INVENTION: Sequence-Directed DNA Binding
TITLE OF INVENTION: Molecules, Compositions and Methods
NUMBER OF SEQUENCES: 664
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genelabs Technologies, Inc.
STREET: 505 Penobscot Drive
CITY: Redwood City
STATE: CA
COUNTRY: USA

Mon Apr 3 08:24:33 2000

us-08-852-495c-1_copy_90000_120000.in1

Page 6

```
ZIP: 94063
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
  FILING DATE: 07-JUN-1995
  APPLICATION NUMBER: US 08/483,080A
  FILING DATE: 20-DEC-1993
  APPLICATION NUMBER: US 08/171,389
  FILING DATE: 17-SEP-1993
  APPLICATION NUMBER: US 08/123,936
  FILING DATE: 23-DEC-1992
  APPLICATION NUMBER: US 07/996,783
  FILING DATE: 27-JUN-1991
  APPLICATION NUMBER: US 07/723,618
  FILING DATE: 22-JUN-1993
  APPLICATION NUMBER: US 08/081,070
  FILING DATE: 22-JUN-1993
  ATTORNEY/AGENT INFORMATION:
    NAME: Brady, John F.
    REGISTRATION NUMBER: 39,118
    REFERENCE/DOCKET NUMBER: 4600-0175.20/G19P3D1
    TELEPHONE: (650) 324-0880
    TELEFAX: (650) 324-0960
  INFORMATION FOR SEQ ID NO: 159:
    SEQUENCE CHARACTERISTICS:
      LENGTH: 48 base pairs
      TYPE: nucleic acid
      STRANDEDNESS: double
      TOPOLOGY: linear
    MOLECULE TYPE: DNA (genomic)
    HYPOTHETICAL: NO
    ORIGINAL SOURCE:
      INDIVIDUAL ISOLATE: Human H4/a gene for H4 histone
      US-08-482-080A-159

Query Match          0.28; Score 48; DB 5; Length 48;
Best Local Similarity 100.0%; Pred. No. 4.1e-07;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22582 GAGAGAAATGGAACCGATCTTTATATCTACGTTACCCCTGCCCCC 22629
DB 48 GAGAGAAATGGAACCGATCTTTATATCTACGTTACCCCTGCCCCC 1

RESULT 13
PCR-US93-12388-159/C
Sequence 159, Application PC/TUS9312388
GENERAL INFORMATION:
  APPLICANT:
  TITLE OF INVENTION: Sequence-Directed DNA Binding
  NUMBER OF SEQUENCES: 641
  CORRESPONDENCE ADDRESSES:
    ADDRESSEE: Genelabs Technologies, Inc.
    STREET: 505 Penobscot Drive
    CITY: Redwood City
    STATE: CA
    COUNTRY: USA
  ZIP: 94063
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
```

```
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: PCT/US93/12388
  FILING DATE:
  CLASSIFICATION:
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: US 08/123,936
    FILING DATE: 17-SEP-1993
    APPLICATION NUMBER: US 07/996,783
    FILING DATE: 23-DEC-1992
  ATTORNEY/AGENT INFORMATION:
    NAME: Fabian GARY R.
    REGISTRATION NUMBER: 33,875
    REFERENCE/DOCKET NUMBER: 4600-0175.41/G19PCT2
    TELECOMMUNICATION INFORMATION:
      TELEPHONE: (415) 324-0880
      TELEFAX: (415) 324-0960
  INFORMATION FOR SEQ ID NO: 159:
    SEQUENCE CHARACTERISTICS:
      LENGTH: 48 base pairs
      TYPE: nucleic acid
      STRANDEDNESS: double
      TOPOLOGY: linear
    MOLECULE TYPE: DNA (genomic)
    HYPOTHETICAL: NO
    ORIGINAL SOURCE:
      INDIVIDUAL ISOLATE: Human H4/a gene for H4 histone
      PCT-US93-12388-159

Query Match          0.28; Score 48; DB 6; Length 48;
Best Local Similarity 100.0%; Pred. No. 4.1e-07;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22582 GAGAGAAATGGAACCGATCTTTATATCTACGTTACCCCTGCCCCC 22629
DB 48 GAGAGAAATGGAACCGATCTTTATATCTACGTTACCCCTGCCCCC 1

RESULT 14
US-08-171-389-170
Sequence 170, Application US/08171389
Patent No. 5578444
GENERAL INFORMATION:
  APPLICANT: Edwards, Cynthia A.
  APPLICANT: Cantor, Charles R.
  APPLICANT: Andrews, Beth M.
  APPLICANT: Tulin, Lisa M.
  APPLICANT: FTY, Kirk E.
  TITLE OF INVENTION: Sequence-Directed DNA Binding
  NUMBER OF SEQUENCES: 641
  CORRESPONDENCE ADDRESSES:
    ADDRESSEE: Genelabs Technologies, Inc.
    STREET: 505 Penobscot Drive
    CITY: Redwood City
    STATE: CA
    COUNTRY: USA
  ZIP: 94063
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US 08/171,389
  FILING DATE:
  CLASSIFICATION: 435
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: US 08/123,936
    FILING DATE: 17-SEP-1993
    APPLICATION NUMBER: US 07/996,783
```

FILING DATE: 23-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/723,618
FILING DATE: 27-JUN-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/081,070
FILING DATE: 22-JUN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Fabian, Gary R.
REGISTRATION NUMBER: 33,875
REFERENCE/DOCKET NUMBER: 4600-0175/G19P3
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 170:
SEQUENCE CHARACTERISTICS:
LENGTH: 47 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: Human histone H3 gene
US-08-171-389-170

Query Match 0.2%; Score 47; DB 1; Length 47;
Best Local Similarity 100.0%; Pred. No. 8.8e-07;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12078 CAAATAGTGGTGTGACTCTATAAAGAGAGTAGCTTCTTCCCT 12124
DB 1 CAAATAGTGGTGTGACTCTATAAAGAGAGTAGCTTCTTCCCT 47

RESULT 15
US-08-123-936-170
Sequence 170, Application US/08123936
Patent No. 5726014
GENERAL INFORMATION:
APPLICANT: Edwards, Cynthia A.
APPLICANT: Cantor, Charles R.
APPLICANT: Andrews, Beth M.
APPLICANT: Turin, Lisa M.
TITLE OF INVENTION: Screening Assay for the Detection of
NUMBER OF SEQUENCES: 640
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genelabs Technologies, Inc.
STREET: 505 Penobscot Drive
CITY: Redwood City
STATE: CA
COUNTRY: USA
ZIP: 94063
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/123,936
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/996,783
FILING DATE: 23-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/723,618
FILING DATE: 27-JUN-1991
ATTORNEY/AGENT INFORMATION:
NAME: Fabian, Gary R.
REGISTRATION NUMBER: 33,875

REFERENCE/DOCKET NUMBER: 4600-0075.32/G19P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 170:
SEQUENCE CHARACTERISTICS:
LENGTH: 47 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: Human histone H3 gene
US-08-123-936-170

Query Match 0.2%; Score 47; DB 2; Length 47;
Best Local Similarity 100.0%; Pred. No. 8.8e-07;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12078 CAAATAGTGGTGTGACTCTATAAAGAGAGTAGCTTCTTCCCT 12124
DB 1 CAAATAGTGGTGTGACTCTATAAAGAGAGTAGCTTCTTCCCT 47

Search completed: March 30, 2000, 17:28:35
Job time: 135577 sec

Mon Apr 3 08:24:33 2000

us-08-852-495c-1_copy_90000_120000.rn1

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 30, 2000, 03:49:55 ; Search time 8198.5 Seconds

(without alignments)
13816.393 Million cell updates/sec

Title: US-08-852-495C-1_COPY_90000_120000

Perfect score: 30001
Sequence: 1 TGCCAGCTCAGTCAGTGGCG.....TATCTAAGCCATGTCAT 30001

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 4538634 seqs, 1887831982 residues

Word size : 0

Total number of hits satisfying chosen parameters: 130662

Minimum DB seq length: 8

Maximum DB seq length: 105

Post-processing: Listing first 45 summaries

Database : ESR:*
1: em_est1:*
2: em_est2:*
3: em_est3:*
4: em_est4:*
5: em_est5:*
6: em_est6:*
7: em_est7:*
8: em_est8:*
9: em_est9:*
10: em_est10:*
11: em_est11:*
12: em_est12:*
13: em_est13:*
14: em_est14:*
15: em_est15:*
16: em_est16:*
17: em_est17:*
18: em_est18:*
19: em_est19:*
20: gb_est1:*
21: gb_est2:*
22: gb_est3:*
23: gb_est4:*
24: gb_est5:*
25: gb_est6:*
26: gb_est7:*
27: gb_est8:*
28: gb_est9:*
29: gb_est10:*
30: gb_est11:*
31: gb_est12:*
32: gb_est13:*
33: gb_est14:*
34: gb_est15:*
35: gb_est16:*
36: gb_est17:*
37: gb_est18:*
38: gb_est19:*
39: gb_est20:*
40: gb_est21:*
41: gb_est22:*
42: gb_est23:*
43: gb_est24:*

44: gb_est25:*
45: gb_est26:*
46: gb_est27:*
47: gb_est28:*
48: gb_est29:*
49: gb_est30:*
50: gb_est31:*
51: gb_est32:*
52: gb_est20:*
53: em_est21:*
54: em_est22:*
55: em_est23:*
56: em_est24:*
57: em_est25:*
58: em_est26:*
59: gb_est33:*
60: gb_est34:*
61: gb_est35:*
62: gb_est36:*
63: gb_est37:*
64: gb_est38:*
65: em_est27:*
66: em_est28:*
67: em_est29:*
68: em_est30:*
69: gb_est39:*
70: gb_est40:*
71: gb_est41:*
72: gb_est42:*
73: gb_est43:*
74: gb_est44:*
75: em_est31:*
76: em_est32:*
77: em_est33:*
78: em_est34:*
79: gb_gss1:*
80: gb_gss2:*
81: gb_gss3:*
82: gb_gss4:*
83: em_gss1:*
84: em_gss2:*
85: em_gss3:*
86: em_gss4:*
87: gb_gss5:*
88: gb_gss6:*
89: gb_gss7:*
90: gb_gss8:*
91: gb_gss9:*
92: em_gss5:*
93: em_gss6:*
94: em_gss7:*
95: em_gss8:*
96: em_gss9:*
97: em_gss10:*
98: em_gss11:*
99: gb_gss10:*
100: gb_gss11:*
101: em_gss12:*
102: gb_gss13:*
103: gb_gss13:*
104: gb_gss14:*
105: gb_gss15:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
C 1	53	0.2	98 31 AA280198	AA280198 2t04b12.r

```

C 3 48 0.2 70 21 T64089
C 4 45 0.1 102 21 A0319270
C 5 43 0.1 91 50 F29374
C 6 42 0.1 95 34 A457423
C 7 41 0.1 92 34 A1932967
C 8 40 0.1 104 105 A062168
C 9 40 0.1 52 20 D25879
C 10 39 0.1 103 81 B48914
C 11 38 0.1 90 61 A1864984
C 12 35 0.1 35 60 A1801185
C 13 35 0.1 88 36 A809831
C 14 35 0.1 96 22 R67088
C 15 35 0.1 102 48 A1591270
C 16 35 0.1 103 104 A0582186
C 17 35 0.1 105 61 A1832832
C 18 34 0.1 104 35 A578614
C 19 34 0.1 100 35 A578614
C 20 34 0.1 57 36 A5634736
C 21 33 0.1 68 28 D4637
C 22 33 0.1 97 25 N49638
C 23 33 0.1 102 81 B48088
C 24 33 0.1 105 28 A084476
C 25 33 0.1 105 30 A247810
C 26 32 0.1 61 50 A1671034
C 27 32 0.1 95 44 A1281558
C 28 32 0.1 95 37 A858120
C 29 32 0.1 103 37 A859571
C 30 32 0.1 103 36 A8807640
C 31 32 0.1 104 22 A128957
C 32 32 0.1 37 22 R70733
C 33 31 0.1 71 35 A4551166
C 34 31 0.1 82 36 A4649287
C 35 31 0.1 90 20 Z25807
C 36 31 0.1 92 40 A5916300
C 37 31 0.1 96 50 A1708433
C 38 31 0.1 98 24 H67549
C 39 31 0.1 101 30 A235077
C 40 31 0.1 101 30 A835205
C 41 31 0.1 102 36 A654562
C 42 30 0.1 34 25 N77004
C 43 30 0.1 60 37 A704865
C 44 30 0.1 81 B36140

```

ALIGNMENTS

```

RESULT 1
AA280198/c 98 bp mRNA EST 14-AUG-1997
DEFINITION z040d12.r1 NCI-CGAP-GCBI Homo sapiens cDNA clone IMAGE:712127 5'
similar to contatans Alu repetitive element; contains element MER22
repetitive element ; mRNA sequence.
ACCESSION AA280198
VERSION AA280198.1 GI:1921755
SOURCE EST.
ORGANISM human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
1 (bases 1 to 98)
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
unpublished (1997)
JOURNAL On Sep 12 1996 this sequence version replaced gi:1395022.
COMMENT Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.

```

```

FEATURES
Source Putative full length read
the vector to vector length is 99
Insert length: 622 Std Error: 0.00
Seq primer: -28m13 rev2 Err from Amersham.
Location/Qualifiers
1..98
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:712127"
/clone_lib="NCI-CGAP-GCBI"
/tissue_type="germinal center B cell"
/lab_host="DH10B"
/Note="Vector: p773D-Pac (Pharmacia) with a modified
polylinker. Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was prepared from human tonsillar cells enriched for
germinal center B cells by flow sorting (CD20+, IgD-),
provided by Dr. Louis M. Staudt (NCI). Dr. David Allman
(NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was
primed with a Not I - oligo(dT) primer
[5'-TGTACCACTGAGAGGAGGAGGCGCCGCAATTTTCTTTT-
3']. Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified p773 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."

```

```

BASE COUNT 25 a 25 c 33 g 15 t
ORIGIN

```

```

Query Match 0.2%; Score 53; DB 31; Length 98;
Best Local 100.00%; Pred. No. 7.3e-08;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2568 CCCAGCTGAGTGCATGCGCGATCTGCGCTCACTGCACCTCGCTTCC 2620
DB 79 CCCAGCTGAGTGCATGCGCGATCTGCGCTCACTGCACCTCGCTTCC 27

```

```

RESULT 2
T64089/c 70 bp mRNA EST 17-FEB-1995
DEFINITION y050f08.r1 Stratiogene lung (#937210) Homo sapiens cDNA clone
IMAGE:7815 5' similar to contatans Alu repetitive element; mRNA
sequence.
ACCESSION T64089
VERSION T64089.1 GI:667954
SOURCE EST.
ORGANISM human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chapell, B.,
Chissole, S., Dietrich, N., Dubuque, T., Favell, A., Gish, W.,
Hawkins, M., Hultman, M., Kucaba, T., Lacy, M., Le, M., Le, N.,
Mardis, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, B.,
Rohlfing, T., Schellenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J.,
Trevaskis, E., Underwood, K., Wohldmann, P., Waterston, R., Wilson, R.
and Marra, M.
Generation and analysis of 280,000 human expressed sequence tags
Genome Res. 6 (9), 807-828 (1996)
97044478
Contact: Willson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.edu
High quality sequence steps: 54
Source: IMAGE Consortium, LNL
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: M13RP1

```

```

JOURNAL On Sep 12 1996 this sequence version replaced gi:1395022.
COMMENT Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.

```


LOCUS AA57423 95 bp mRNA EST 06-JUN-1997
 DEFINITION aa5602.r1 Stratagene fetal retina 937202 Homo sapiens cDNA clone
 IMAGE:838155 5' similar to contains Alu repetitive element; mRNA
 sequence.

ACCESSION AA57423
 VERSION AA57423.1 GI:2180143
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 95)
 Haller, L., Allen, M., Bowles, L., Dubuque, T., Gelsel, G., Jost, S.,
 Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J.,
 Moore, B., Schellenberg, R., Steptoe, M., Tan, F., Theisling, B.,
 White, Y., Wylie, T., Waterston, R. and Wilson, R.
 WashU-Merck EST Project 1997
 Unpublished (1997)
 On Nov 4, 1996 this sequence version replaced gi:1671271.

TITLE JOURNAL
 COMMENT Contact: Wilson RK
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@wustl.wustl.edu
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Seq primer: -28ml3 rev1 bt from Amersham
 High quality sequence stop: 76.

FEATURES
 source
 Location/Qualifiers
 1..95
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:838155"
 /clone_lib="Stratagene fetal retina 937202"
 /sex="mixed"
 /lab_host="SOLR (kanamycin resistant)"
 /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
 XhoI; Cloned unidirectionally. Primer: Oligo dT. Pooled
 retinal tissue. Average insert size: 1.0 kb; Uni-ZAP XR
 Vector: -5' adaptor sequence: 5' GAATTCGCGACG 3' -3'
 adaptor sequence: 5' CTCGAGTCTTTTCTTTTCTTTT 3' "

BASE COUNT 14 a 39 c 25 g 17 t

ORIGIN
 Query Match 0.1%; Score 42; DB 34; Length 95;
 Best Local Similarity 100.0%; Pred. No. 0.00046;
 Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1294 TCCGAGCTACTCGGAGCTGAGCGAGAGATGCGGTAC 1335
 |||||||
 Db 84 TCCGAGCTACTCGGAGCTGAGCGAGAGATGCGGTAC 43

RESULT 6
 A1932967/c 92 bp mRNA EST 02-SEP-1999
 LOCUS w040f09.x1 NCI_CGAP_Gas4 Homo sapiens cDNA clone IMAGE:2457833 3'
 DEFINITION mRNA sequence.
 A1932967
 A1932967.1 GI:5671704
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 92)
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 On Dec 20, 1995 this sequence version replaced gi:1134104.

JOURNAL
 COMMENT

CONTACT: Robert Strausberg, Ph.D.
 Tel: (301) 496-1550
 Email: Robert.Strausberg@nih.gov
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
 Emmert-Buck, M.D., Ph.D.
 CDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 www.bio.llnl.gov/dbp/image/image.html

Seq primer: -40UP from Gibco.

FEATURES
 source
 Location/Qualifiers
 1..92
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:2457833"
 /clone_lib="NCI_CGAP_Gas4"
 /tissue_type="poorly differentiated adenocarcinoma with
 signed ring cell features"
 /lab_host="DH10B"
 /note="Organ: stomach; Vector: pCMV-SPORT6; Site_1: SalI;
 Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
 Average insert size 1.69 kb. Life Technologies catalog #: 11549-011"

BASE COUNT 29 a 19 c 14 g 30 t

ORIGIN
 Query Match 0.1%; Score 41; DB 62; Length 92;
 Best Local Similarity 100.0%; Pred. No. 0.001;
 Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1392 GACAGAGCGAGACTCGCTCTCAAAAAAAAAAAAAAAAA 1432
 |||||||
 Db 41 GACAGAGCGAGACTCGCTCTCAAAAAAAAAAAAAAAAA 1

RESULT 7
 A0626158/c 104 bp DNA GSS 16-JUN-1999
 LOCUS CITBI-El-2650P18.TR CITBI-El Homo sapiens genomic clone 2650P18,
 DEFINITION genomic survey sequence.
 A0626158
 A0626158.1 GI:5088550
 KEYWORDS GSS.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 104)
 Zhao, S., Adams, M.D., Nierman, W., Malek, J., Shizuya, H., Simon, M. and
 Venter, J.C.
 Use of BAC End Sequences from Caltech Libraries for Sequence-Ready
 Map Building
 Unpublished (1997)
 Other_GSSs: CITBI-El-2650P18.TF
 Contact: Shaying Zhao, William Nierman, Mark Adams
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: hbe@igrr.org
 Clones are available from Research Genetics (info@resgen.com). BAC
 end search Page:
 http://www.tigr.org/tdb/hunggen/bac_end_search/bac_end_search.html.
 Seq primer: M13 Reverse
 Class: BAC ends

FEATURES
 source
 Location/Qualifiers
 1..104
 /organism="Homo sapiens"

/db_xref="taxon:9606"
/clone="2650P18"
/clone_id="CTBI-E1"
/sex="male"
/cell_type="sperm"
/note="Vector: pBelOAC11; Site_1: EcoRI; Site_2: EcoRI;
Caltech Human BAC Library P"
BASE COUNT 12 a 28 c 22 g 42 t
ORIGIN

Query Match 0.1%; Score 41; DB 105; Length 104;
Best Local Similarity 100.0%; Pred. No. 0.00098;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1392 GACAGAGCAGAGCTCGTCAAAAAAAAAAAAAAAAAAAAA 1432
|||||
Db 94 GACAGAGCAGAGCTCGTCAAAAAAAAAAAAAAAAAAAAA 54

RESULT 8
D25879 52 bp mRNA EST 30-NOV-1995
LOCUS HMGS05672 Human colon mucosa Homo sapiens cDNA clone cm2335 3',
DEFINITION mRNA sequence.
ACCESSION D25879
VERSION D25879.1 GI:500543
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 52)
AUTHORS Okubo,K., Yoshii,J., Yokouchi,H., Kameyama,M. and Matsubara,K.
TITLE Global analysis of gene expression in colon mucosa: a large scale
random cDNA sequencing analysis
JOURNAL Contact: Okubo,K., Itoh,K., Yoshii,J., Yokouchi,H. and Matsubara,K.
COMMENT Institute for Molecular and Cellular Biology
Osaka University
3-1 Yamada-oka, Suita, Osaka 565, Japan.
FEATURES
SOURCE Location/Qualifiers
1..52
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="cm2335"
/clone_id="Human colon mucosa"
/note="Adult male, tissue type = colon mucosa"
BASE COUNT 12 a 17 c 14 g 9 t
ORIGIN

Query Match 0.1%; Score 40; DB 20; Length 52;
Best Local Similarity 100.0%; Pred. No. 0.0031;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1376 GCACCTCAGCTGGGTGACAGAGCAGACTCGTCTCAA 1415
|||||
Db 13 GCACCTCAGCTGGGTGACAGAGCAGACTCGTCTCAA 52

RESULT 9
B48914 103 bp DNA GSS 08-APR-1999
LOCUS RPI11-4A12.TP RPI1-11 Homo sapiens genomic clone RPI1-11-4A12,
DEFINITION genomic survey sequence.
ACCESSION B48914
VERSION B48914.1 GI:2601151
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE 1 (bases 1 to 103)
AUTHORS Adams,M.D., Rounsley,S.D., Field,C.E., Bass,S., Linher,K.,
Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., de Jong,P. and
Venter,J.C.
TITLE Use of BAC End Sequences for Sequence-Ready Map Building
JOURNAL Unpublished (1997)
COMMENT Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Clones are derived from the human BAC library RPI1-11. For BAC
library availability, please contact Pieter de Jong
(pieter@dejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (<http://bacpac.med.buffalo.edu/ordering>) or from
Research Genetics (info@resgen.com). BAC end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
Seq primer: SP6
Class: BAC ends.
FEATURES
SOURCE Location/Qualifiers
1..103
/organism="Homo sapiens"
/db_xref="GDB:750163"
/db_xref="taxon:9606"
/clone="RPI1-11-4A12"
/clone_id="RPI1-11"
/sex="Male"
/cell_type="Lymphocytes"
/note="Vector: pBAC3.6; Site_1: EcoRI; Site_2: EcoRI;
RPI11 Human Male BAC Library"
BASE COUNT 30 a 28 c 30 g 15 t
ORIGIN

Query Match 0.1%; Score 40; DB 81; Length 103;
Best Local Similarity 100.0%; Pred. No. 0.0022;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1168 TTGGAGAGCCGAGCGCGGATCAGAGTCAAGAGAT 1207
|||||
Db 14 TTGGAGAGCCGAGCGCGGATCAGAGTCAAGAGAT 53

RESULT 10
A1864984 90 bp mRNA EST 30-AUG-1999
LOCUS WK06c11.X1 NCI_CGAP_Lym12 Homo sapiens cDNA clone IMAGE:2411540 3',
DEFINITION similar to contains A18 repetitive element; mRNA sequence.
ACCESSION A1864984
VERSION A1864984.1 GI:5529091
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 90)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT On May 18, 1998 this sequence version replaced g1:3137794.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Life Technologies catalog #: 11547-015
DNA sequencing by: Washington University Genome Sequencing Center
clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LINL at:
www-bio.llnl.gov/bdrip/image/image.html
Seq primer: -40UP from G1bco.

Mon Apr 3 08:24:34 2000

us-08-852-495c-1-copy_90000_120000.1st

Page 6

FEATURES
source
1. 90
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="241540"
/clone_id="NCI-CGAP-Lym12"
/tissue_type="lymphoma, follicular mixed small and large cell"
/lab_host="DH10B"
/note="Organ: lymph node; Vector: PCMV-SPORT6; Site_1: SalI; Site_2: NotI; Cloned unidirectionally. Primer: 0190 dt. Average insert size 1.25 kb. Life Technologies catalog #: 11547-015" 21 t

BASE COUNT
13 a 29 c 27 g 21 t

ORIGIN

Query Match
Best Local Similarity 100.0%; Pred. No. 0.025; Length 90;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1337 CGGAGGCGGAGCTGCGAGTGGCGGAGTCCGCCA 1373

DB 79 CGGAGCGGAGCTGCGAGTGGCGGAGTCCGCCA 43

RESULT 11
LOCUS
DEFINITION
AI801185 35 bp mRNA EST 06-JUL-1999
to79h04.x1 NCI-CGAP_Gas4 Homo sapiens cDNA clone IMAGE:2184535 3'
similar to gb|U87921|HUMALCD120 Human carcinoma cell-derived Alu
RNA transcript, (rRNA) contains element TARI repetitive element ;,
RNA sequence.
AI801185
AI801185.1 GI:5366657
EST.
SOURCE
ORGANISM
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 35)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
On Mar 16 1998 this sequence version replaced gi:2961671.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Christopher Koskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LNM at:
www.bio.lnl.gov/bdrrp/image/image.html

JOURNAL
COMMENT
Unpublished (1997)
On Mar 16 1998 this sequence version replaced gi:2961671.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Christopher Koskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LNM at:
www.bio.lnl.gov/bdrrp/image/image.html

FEATURES
source
1. 35
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="2184535"
/clone_id="NCI-CGAP_Gas4"
/tissue_type="poorly differentiated adenocarcinoma with
signet ring cell features"
/lab_host="DH10B"
/note="Organ: stomach; Vector: PCMV-SPORT6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: 0190 dt.
Average insert size 1.69 kb. Life Technologies catalog #:
11549-011"

BASE COUNT
5 a 10 c 10 g 10 t

ORIGIN

Query Match
Best Local Similarity 100.0%; Pred. No. 0.2; Length 35;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1382 CAGCTGGGTGACAGAGCGAGCTCGCTCAAAA 1416

DB 35 CAGCTGGGTGACAGAGCGAGCTCGCTCAAAA 1

RESULT 12
LOCUS
DEFINITION
AA809831 88 bp mRNA EST 19-FEB-1998
oa40fil.s1 NCI-CGAP_GCB1 Homo sapiens cDNA clone IMAGE:1307469 3'
similar to contains Alu repetitive element; , RNA sequence.
AA809831
AA809831.1 GI:2879237
EST.
SOURCE
ORGANISM
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 88)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
On Jan 19 1998 this sequence version replaced gi:2287025.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Louis M. Staudt, M.D., David Allman,
Ph.D., Gerald Marti, M.D.
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LNM at:
www.bio.lnl.gov/bdrrp/image/image.html

JOURNAL
COMMENT
Insert Length: 963 Std Error: 0.00
Seq primer: -40m13 fwd. pm from Amersham
High quality sequence stop: 50.
Location/Qualifiers
1. 88
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="1307469"
/clone_id="NCI-CGAP_GCB1"
/tissue_type="germinal center B cell"
/lab_host="DH10B"
/note="Vector: pTZ19-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand CDNA
was prepared from human tonsillar cells enriched for
germinal center B cells by flow sorting (CD20+, 19D-),
provided by Dr. Louis M. Staudt (NCI), Dr. David Allman
(NCI) and Dr. Gerald Marti (CBER). CDNA synthesis was
primed with a Not I - 0190(dt) primer
15'-TGTACCAATCGAAGTGGAGCGCGCCGCTCATTTTCTTTTCTTTT-
3/1. Double-stranded CDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pTZ19 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT
14 a 28 c 23 g 23 t

ORIGIN

Query Match
0.1%; Score 35; DB 38; Length 88;

REFERENCE
AUTHORS
TITLE
COMMENT

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominoidea; Homo.
1 (bases 1 to 103)
Zhao, S., Adams, M.D., Niernan, W., Malek, J., de Jong, P. and
Venter, J.C.
Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
Map Building
Unpublished (1997)
Contract: Shaying Zhao, William Niernan, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208
Email: hbe@tigr.org

Clones are derived from the human BAC library RPCI-11. For BAC
library availability, please contact Pieter de Jong
(pieter@jng.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (<http://bacpac.med.buffalo.edu/ordering>) or from
Research Genet cs (http://www.tigr.org/tdb/hungen/bac_end_search/bac_end_search.html).
Seq primer: T7
Class: BAC ends.

FEATURES
SOURCE

location/Qualifiers
1..103
/organism="Homo sapiens"
/db_xref="GeneID:7621533"
/db_xref="taxon:9606"
/clone="RPCI-11-317H22"
/clone_lib="RPCI-11"
/sex="Male"
/cell_type="Lymphocytes"
/note="Vector: pBAC3.6; Site_1: EcoRI; Site_2: EcoRI;
RPCI11 Human Male BAC Library"
BASE COUNT 31 a 27 c 27 g 18 t
ORIGIN

Query Match 0.1%; Score 35; DB 104; Length 103;
Best Local Similarity 100.0%; Pred. No. 0.12;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 5307 CCAGCTGTCTGACTGCTGACCTCAAGTGATC 5341
DB 63 CCAGCTGTCTGACTGCTGACCTCAAGTGATC 29

Search completed: March 30, 2000, 13:20:07
Job time: 122661 sec



1
2
3

1. The first part of the document discusses the importance of maintaining accurate records of all transactions. It emphasizes that proper record-keeping is essential for ensuring the integrity of the financial data and for facilitating the audit process. The document also highlights the need for transparency and accountability in all financial reporting.

2. The second part of the document outlines the specific procedures for recording and reconciling transactions. It provides a detailed description of the steps involved in the accounting cycle, from identifying the transaction to posting the entry to the ledger and finally to the preparation of the financial statements. The document also includes a discussion of the various types of accounts and how they are classified.

3. The third part of the document discusses the importance of internal controls in preventing fraud and errors. It describes the various types of internal controls, such as segregation of duties, authorization, and physical controls, and explains how they are implemented in the accounting system. The document also includes a discussion of the role of the internal auditor in monitoring and evaluating the effectiveness of the internal controls.

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 30, 2000, 17:03:40 ; Search time 13413.1 Seconds

(without alignments)
-6791.452 Million cell updates/sec

Title: US-08-852-495C-1_COPY_115000_145000

Perfect score: 30001

Sequence: 1 AAATCTTCATTTTGTATA.....AGAAATTAACCTGATTAAA 30001

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 821193 seqs, -1518192014 residues

Word size : 0

Total number of hits satisfying chosen parameters: 345848

Minimum DB seq length: 8

Maximum DB seq length: 105

Post-processing: Listing first 45 summaries

Database : GenEmb1:*
1: gb_ba1:*
2: gb_ba2:*
3: gb_ov:*
4: gb_ov:*
5: gb_ov:*
6: gb_ov:*
7: gb_ov:*
8: gb_ov:*
9: gb_ov:*
10: gb_ov:*
11: gb_ov:*
12: gb_ov:*
13: gb_ov:*
14: gb_ov:*
15: gb_ov:*
16: gb_ov:*
17: gb_ov:*
18: gb_ov:*
19: gb_ov:*
20: gb_ov:*
21: gb_ov:*
22: gb_ov:*
23: gb_ov:*
24: gb_ov:*
25: gb_ov:*
26: gb_ov:*
27: gb_ov:*
28: gb_ov:*
29: gb_ov:*
30: gb_ov:*
31: gb_ov:*
32: gb_ov:*
33: gb_ov:*
34: gb_ov:*
35: gb_ov:*
36: gb_ov:*
37: gb_ov:*
38: gb_ov:*
39: gb_ov:*
40: gb_ov:*
41: gb_ov:*
42: gb_ov:*
43: gb_ov:*

44: gb_htg6:*
45: gb_htg7:*
46: em_htg1:*
47: em_htg2:*
48: em_htg3:*
49: em_htg4:*
50: gb_pl3:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	38	0.1	60	5 ARO51487	ARO51487 Sequence
2	38	0.1	76	5 ARO51499	ARO51499 Sequence
3	36	0.1	67	9 HUMALUACD	L36843 Homo sapien
4	35	0.1	51	9 HUMALUACD	D17279 Human HepG2
5	35	0.1	79	13 HSU24138	U24138 Hylobates s
6	34	0.1	66	40 AF087511	AF087511 Homo sapi
7	33	0.1	79	10 S73203	S73203 ATL-1 (rand
8	32	0.1	80	9 HUMBRKFAA	M36131 Human alpha
9	32	0.1	80	9 HUMBRKFAA	M36132 Human alpha
10	31	0.1	60	11 HSCBPS115	AF084969 Homo sapi
11	31	0.1	76	11 AF032238	AF032238 Otolomur
12	31	0.1	97	9 HUMDLRA2	M14180 Human low d
13	31	0.1	97	9 HUMDLRA2	M14179 Human fam1
14	31	0.1	103	9 HUMALCE221	M87896 Human fam1
15	30	0.1	80	9 HUMBRKFAA	M36131 Human alpha
16	30	0.1	80	9 HUMBRKFAA	M36132 Human alpha
17	30	0.1	95	9 HSTHFKTB	X66361 H.sapiens m
18	30	0.1	100	9 HSLAS27	X91545 H.sapiens D
19	30	0.1	101	10 S79560	S79560 HRX (intron
20	30	0.1	90	9 HUMDLRFL	K03555 Human low d
21	29	0.1	102	13 G32906	G32906 A009W09 Hum
22	29	0.1	102	13 G37854	G37854 P12689 Plas
23	29	0.1	104	13 G20851	G20851 human STS A
24	28	0.1	50	5 I23510	I23510 Sequence 5
25	28	0.1	50	5 I28359	I28359 Sequence 5
26	28	0.1	50	5 I28514	I28514 Sequence 5
27	28	0.1	50	5 I41125	I41125 Sequence 5
28	28	0.1	50	5 I49056	I49056 Sequence 5
29	28	0.1	50	5 I70295	I70295 Sequence 5
30	28	0.1	50	5 I90068	I90068 Sequence 5
31	28	0.1	57	5 A45333	A45333 Sequence 3
32	28	0.1	57	5 ARO61138	ARO61138 Sequence
33	28	0.1	57	9 HSDNAPYTB	X74432 H.sapiens (
34	28	0.1	50	9 HUMDLRPM	K03556 Human low d
35	28	0.1	103	13 HSBIC8R	X57789 Human sequ
36	28	0.1	40	5 A68621	A68621 Sequence 1
37	27	0.1	63	13 HUMUT5302A	L30829 Human STS U
38	27	0.1	67	9 HUMALUACD	L36843 Homo sapien
39	27	0.1	71	13 HUMUT41A	L28838 Human STS U
40	27	0.1	72	13 HUMUT41A	L30919 Human STS U
41	27	0.1	73	5 A08915	A08915 H.sapiens (
42	27	0.1	80	9 HUMBRKFAA	M36133 Human alpha
43	27	0.1	80	34 DRORS1584	M30167 D.melanogas
44	27	0.1	80	34 DRORS1584	M31539 D.melanogas
45	27	0.1	80	34 DRORS1586	

ALIGNMENTS

RESULT 1
LOCUS ARO51487 60 bp DNA
DEFINITION Sequence 57 from patent US 5830670.
ACCESSION ARO51487
VERSION ARO51487.1 GI:5974851 PAT 29-SEP-1999

KEYWORDS : Unknown
SOURCE : Unknown
ORGANISM : Unclassified

REFERENCE : 1 (bases 1 to 60)
AUTHORS : de la Monte,S. and Wands,J.R.
TITLE : Neural thread protein gene expression and detection of Alzheimer's disease

JOURNAL : Patent: US 5830670-A 57 03-NOV-1998;
FEATURES : Location/Qualifiers
SOURCE : 1. 60
BASE COUNT : 12 a 14 c 15 g 19 t
ORIGIN

Query Match : 0.1%; Score 38; DB 5; Length 60;
Best Local Similarity 100.0%; Pred. No. 2.2e-07;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 17233 TTTTGTATTTTAGTAGAGACAGGTTCCACCGTGT 17269
|||||
Db 16 TTTTGTATTTTAGTAGAGACAGGTTCCACCGTGT 53

RESULT 2
LOCUS : AR051499 76 bp DNA
DEFINITION : Sequence 69 from patent US 5830670.
ACCESSION : AR051499
VERSION : AR051499.1 GI:5974863
KEYWORDS :
SOURCE : Unknown.
ORGANISM : Unknown.
REFERENCE : 1 (bases 1 to 76)
AUTHORS : de la Monte,S. and Wands,J.R.
TITLE : Neural thread protein gene expression and detection of Alzheimer's disease

JOURNAL : Patent: US 5830670-A 69 03-NOV-1998;
FEATURES : Location/Qualifiers
SOURCE : 1. 76
BASE COUNT : 15 a 19 c 18 g 24 t
ORIGIN

Query Match : 0.1%; Score 38; DB 5; Length 76;
Best Local Similarity 100.0%; Pred. No. 2.2e-07;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 17233 TTTTGTATTTTAGTAGAGACAGGTTCCACCGTGT 17269
|||||
Db 16 TTTTGTATTTTAGTAGAGACAGGTTCCACCGTGT 53

RESULT 3
LOCUS : HUMALUAND 67 bp DNA
DEFINITION : Homo sapiens 4000 year old remains from Nekht-anhk Alu repeat
fragment 12:5.
ACCESSION : L36843
VERSION : L36843.1 GI:556196
KEYWORDS :
SOURCE : Alu repeat.
ORGANISM : Homo sapiens (individual isolate 4000 year old remains from Nekht-anhk) liver DNA.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (sites)
REFERENCE :
AUTHORS : Paabo,S.
TITLE : Ancient DNA: extraction, characterization, molecular cloning, and enzymatic amplification

JOURNAL : Proc. Natl. Acad. Sci. U.S.A. 86 (6), 1939-1943 (1989)
MEDLINE : 89184542
FEATURES : Location/Qualifiers
SOURCE : 1. 67
/organism="Homo sapiens"
/isolate="4000 year old remains from Nekht-anhk"
/db_xref="taxon:9606"
/tissue_type="liver"
/rpl_family="Alu"
repeat_region
BASE COUNT : 15 a 18 c 23 g 11 t
ORIGIN

Query Match : 0.1%; Score 36; DB 9; Length 67;
Best Local Similarity 100.0%; Pred. No. 2.1e-06;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 22232 GTGGCTCACACCTGTAATCCACAGCTTGGAGGC 22267
|||||
Db 9 GTGGCTCACACCTGTAATCCACAGCTTGGAGGC 44

RESULT 4
LOCUS : HUMD6803M3/C 51 bp mRNA
DEFINITION : Human HepG2 3' region MBOI cDNA, clone hmd6803m3.
ACCESSION : D17279
VERSION : D17279.1 GI:598922
KEYWORDS :
SOURCE : gene signature
ORGANISM : Homo sapiens Male cell_line:HepG2 cDNA to mRNA, clone_11b:kiserv.
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 51)
REFERENCE :
AUTHORS : Matoba,R.
TITLE : Direct Submission
JOURNAL : Submitted (21-JUL-1993) to the DDBJ/EMBL/GenBank databases. Ryo Matoba, Osaka University, Institute for Molecular and Cellular Bio; 1-3, Yamada-oka, Suita, Osaka 565, Japan (E-mail:matoba@inherit.imcb.osaka-u.ac.jp, Tel:81-6-877-5111(ex.3314), Fax:81-6-877-1922)

REFERENCE :
AUTHORS : Matoba,R., Okubo,K., Hori,N., Fukushima,A. and Matsubara,K.
TITLE : Improves analysis of 5'-coding information to a 3'-directed cDNA library
JOURNAL : Gene 146 (2), 199-207 (1994)
MEDLINE : 94357437
COMMENT : Submitted (21-Jul-1993) to DDBJ by:
Ryo Matoba
Research Institute of Innovative Technology for the Earth 9-2
Kisugawada Kisu-cho,
Soraku-gun, Kyoto
Japan, 619-02
Phone: 07747-5-2308
Fax: 07747-5-2321.
FEATURES : Location/Qualifiers
SOURCE : 1. 51
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_line="HepG2"
/clone_11b="Kiserv"
/sex="Male"
BASE COUNT : 13 a 19 c 12 g 7 t
ORIGIN

Query Match : 0.1%; Score 35; DB 9; Length 51;
Best Local Similarity 100.0%; Pred. No. 6.8e-06;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 16302 CTCTGTCGCCAGGCTGAGTGCAGTGCGGTATC 16336

Db 35 CTTCTGCCAGCGTGCATGCTGCTATC 1
|||||
RESULT 5
LOCUS HSU24138 79 bp DNA STS 08-NOV-1995
DEFINITION Hylobates sp. sequence homologous to Human dinucleotide repeat
region clone Mfd138, sequence tagged site.
ACCESSION U24138
VERSION U24138.1 GI:773649
KEYWORDS STS.
SOURCE gibbon.
ORGANISM Hylobates sp.
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hylobatidae; Hylobates.
1 (bases 1 to 79)
TITLE Arcot,S.S., Wang,Z., Weber,J.L., Deininger,P.L. and Batzer,M.A.
JOURNAL Alu repeats: a source for the genesis of primate microsatellites
MEDLINE Genomics 29 (1), 136-144 (1995)
96079101
2 (bases 1 to 79)
Baker,M.A., Rubin,C.M., Hellmann-Blumberg,U., Alegria-Hartman,M.,
Leeftang,E.P., Stern,J.D., Bazan,H.A., Shaikh,T.H., Deininger,P.L.
and Schmid,C.W.
dispersion and insertion polymorphism in two small subfamilies of
recently amplified human Alu repeats
J. Mol. Biol. 247 (3), 418-427 (1995)
95230683
3 (bases 1 to 79)
Weber,J.L.
REFERENCE Direct Submission
AUTHORS Submitted (30-MAR-1995) James L. Weber, Medical Genetics,
JOURNAL Marshfield Medical Research Foundation, 1000 N. Oak Ave.,
Marshfield, WI 54449, USA
FEATURES
source
1. 79
/organism="Hylobates sp."
/db_xref="taxon:9581"
/chromosome="7"
/note="Human Mfd158 homolog, GenBank Accession Number
L15398"
BASE COUNT 30 a 17 c 15 g 17 t
ORIGIN
Query Match 0.1%; Score 35; DB 13; Length 79;
Best Local Similarity 100.0%; Pred. No. 6.7e-06;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 22361 GCGGTGTCGATGCTGCTATCCACGCTACTC 22395
|||||
Db 45 GCGGTGTCGATGCTGCTATCCACGCTACTC 79
|||||
RESULT 6
LOCUS AF087511 66 bp mRNA PRI 13-SEP-1999
DEFINITION Homo sapiens clone ENAC+22 epithelial sodium channel alpha subunit
(SCNN1A) mRNA, alternatively spliced, partial sequence.
ACCESSION AF087511
VERSION AF087511.1 GI:5870626
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 66)
TITLE Oh,Y. and Warnock,D.G.
JOURNAL An Alu cassette in the human epithelial sodium channel
UNPUBLISHED
REFERENCE 2 (bases 1 to 66)
AUTHORS Oh,Y. and Warnock,D.G.

TITLE Direct Submission
JOURNAL Submitted (26-AUG-1998) Medicine, UAB, 1720 7th Ave. So.,
Birmingham, AL 35294, USA
FEATURES
source
1. 66
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="12"
/clone="ENAC+22"
misc-feature
1. .>66
/gene="SCNN1A"
/note="alternatively spliced epithelial sodium channel
alpha subunit containing intronic Alu sequence"
1. .>66
/gene="SCNN1A"
BASE COUNT 12 a 20 c 24 g 10 t
ORIGIN
gene
Query Match 0.1%; Score 34; DB 40; Length 66;
Best Local Similarity 100.0%; Pred. No. 2.1e-05;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 22235 GCTCACACCTGTATCCACGACTTGGAGGCC 22268
|||||
Db 14 GCTCACACCTGTATCCACGACTTGGAGGCC 47
|||||
RESULT 7
LOCUS S73203 79 bp DNA PRI 28-FEB-1995
DEFINITION ALL-1 (tandem duplication) [human, acute myeloid leukemia patient,
Genomic Mutant, 79 nt].
ACCESSION S73203
VERSION S73203.1 GI:685048
KEYWORDS human acute myeloid leukemia patient.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 79)
REFERENCE Schichman,S.A., Calligaris,M.A., Strout,M.P., Carter,S.L., Gu,Y.,
JOURNAL Cnaan,I., Bloomfield,C.D. and Croce,C.M.
MEDLINE ALL-1 tandem duplication in acute myeloid leukemia with a normal
94320053
Cancer Res. 54 (16), 4277-4280 (1994)
REMARK karyotype involves homologous recombination between Alu elements
Genbank staff at the National Library of Medicine created this
entry [NCBI gblseq 155288] from the original journal article.
This sequence comes from Fig. 3a.
Map location: 11.
FEATURES
source
1. 79
/organism="Homo sapiens"
/db_xref="taxon:9606"
1. 79
/partial
/gene="ALL-1"
BASE COUNT 19 a 18 c 29 g 13 t
ORIGIN
Query Match 0.1%; Score 33; DB 10; Length 79;
Best Local Similarity 100.0%; Pred. No. 6.6e-05;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 22418 GCTTGAACCCAGAGGCGAGGTGAGTGCAGTGCAGC 22450
|||||
Db 37 GCTTGAACCCAGAGGCGAGGTGAGTGCAGTGCAGC 69
|||||
RESULT 8
HUMBRKFAA/c

LOCUS HUMBRKFA 80 bp DNA PRI 27-APR-1993
DEFINITION Human alpha-galactosidase breakpoint region.
ACCESSION M36131.1 GI:179541
VERSION M36131.1
KEYWORDS Fabry disease; breakpoint junction; glycosphingolipid catabolism.
SOURCE Human DNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Kornreich, R., Bishop, D.F. and Desnick, R.J.
TITLE Alpha-galactosidase A gene rearrangements causing Fabry disease:
identification of short direct repeats at breakpoints in an
Alu-rich gene
JOURNAL J. Biol. Chem. 265, 9319-9326 (1990)
MEDLINE 90264427
FEATURES
Location/Qualifiers
1..80
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 19 a 28 c 20 g 13 t
ORIGIN

Query Match 0.1%; Score 32; DB 9; Length 80;
Best Local Similarity 100.0%; Pred. No. 0.00021;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22242 CCTGTAATCCAGCACTTTGGAGGCCAAGT 22273
DB 47 CCTGTAATCCAGCACTTTGGAGGCCAAGT 16

RESULT 9
LOCUS HUMBRKFA 80 bp DNA PRI 27-APR-1993
DEFINITION Human alpha-galactosidase breakpoint family F.
ACCESSION M36132
VERSION M36132.1 GI:179542
KEYWORDS Fabry disease; breakpoint junction; glycosphingolipid catabolism.
SOURCE Human DNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Kornreich, R., Bishop, D.F. and Desnick, R.J.
TITLE Alpha-galactosidase A gene rearrangements causing Fabry disease:
identification of short direct repeats at breakpoints in an
Alu-rich gene
JOURNAL J. Biol. Chem. 265, 9319-9326 (1990)
MEDLINE 90264427
FEATURES
Location/Qualifiers
1..80
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 17 a 29 c 16 g 18 t
ORIGIN

Query Match 0.1%; Score 32; DB 9; Length 80;
Best Local Similarity 100.0%; Pred. No. 0.00021;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22242 CCTGTAATCCAGCACTTTGGAGGCCAAGT 22273
DB 47 CCTGTAATCCAGCACTTTGGAGGCCAAGT 16

RESULT 10
LOCUS HSCBRB515/c 60 bp DNA PRI 26-SEP-1998
DEFINITION Homo sapiens core-binding factor beta subunit (CBFB) gene, intron
5, partial sequence.

ACCESSION AF084969
VERSION AF084969.1 GI:3659831
KEYWORDS 15 of 16
SEGMENT
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS van der Reijden, B.A., Dauwerse, H.G., Gilles, R.H.,
Jagmoohan-Changur, S., Wijnga, C., Liu, P.P., Smit, B., Wessels, H.W.,
Beverloot, G.C., Jottrand-Bellomo, M., Martinet, D.,
Chlematier, D.M.F., Lafage-Pochitaloff, M., Gabert, J., Reiffers, J.,
Bilhou-Nabera, C.E., van Ommen, G.-J.B., Hagemeijer, A. and
Breuning, M.H.
TITLE Genomic acute myeloid leukemia-associated inv(16)(p13q22)
breakpoints are tightly clustered
JOURNAL Oncogene (1998) in press
REFERENCE
AUTHORS van der Reijden, B.A., Dauwerse, H.G., Gilles, R.H.,
Jagmoohan-Changur, S. and Breuning, M.H.
TITLE Nonoverlapping genomic inv(16)(p13q22) CBFB intron 5 breakpoint
regions ordered centromeric to telomeric
JOURNAL unpublished
AUTHORS van der Reijden, B.A., Dauwerse, H.G., Gilles, R.H.,
Jagmoohan-Changur, S. and Breuning, M.H.
TITLE Direct Submission
JOURNAL Submitted (20-AUG-1998) Hematology, Erasmus University Rotterdam,
Dr. Molwaterplein 50, Rotterdam 3015 GE, The Netherlands
FEATURES
Location/Qualifiers
1..60
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="16"
/map="16q22"
BASE COUNT 13 a 20 c 15 g 12 t
ORIGIN

Query Match 0.1%; Score 31; DB 11; Length 60;
Best Local Similarity 100.0%; Pred. No. 0.00066;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 23887 TCATGCGTATCCAGCACTTTGGAGGC 23917
DB 39 TCATGCGTATCCAGCACTTTGGAGGC 9

RESULT 11
LOCUS AF032238/c 76 bp DNA PRI 13-OCT-1998
DEFINITION Otlemur crassicaudatus clone GALLI72 Galago Alu.
ACCESSION AF032238
VERSION AF032238.1 GI:2642069
KEYWORDS
SOURCE
ORGANISM Otlemur crassicaudatus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Strepsirhini; Galagonidae; Otlemur.
REFERENCE
AUTHORS Zietkiewicz, E., Richer, C., Smet, D. and Labuda, D.
TITLE Monophyletic origin of Alu elements in primates
JOURNAL J. Mol. Evol. 47 (2), 172-182 (1998)
MEDLINE 98360010
AUTHORS Zietkiewicz, E., Richer, C., Smet, D. and Labuda, D.
TITLE Direct Submission
JOURNAL Submitted (31-OCT-1997) Pediatrics, University of Montreal,
Sainte-Justine Hospital, Research Center (Charles Bruneau Center of
Cancerology), 3175 Cote Sainte-Catherine, Montreal, Quebec H3T-1C5,
Canada
COMMENT Sequences corresponding to the PCR primers used to amplify

Alu element are not included.

Forward primer: 5'-ggcgcggtggtccacg-3';

Reverse primer: 5'-tttttgagcagcgagtcgtcc-3'.

FEATURES
source
Location/Qualifiers
1..76
/organism="Otolemur crassicaudatus"
/db_xref="taxon:9463"
/clone="GAL1127"
/note="PCR-amplified"

repeat_region
1..76
/rpt_family="Galago Alu"
/rpt_type="dispersed"

BASE COUNT
20 a 20 c 21 g 15 t
ORIGIN

Query Match
Best Local Similarity 100.0%; Score 31; DB 11; Length 76;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7702 TGGCTCAAACTCCTGAGCTCAAGCAATCCA 7732
|||||
Db 66 TGGCTCAAACTCCTGAGCTCAAGCAATCCA 36

RESULT 12
HUMDLRA2/c
LOCUS
DEFINITION Human low density lipoprotein receptor gene (LDLR), intron 5
(partial).
ACCESSION M14180.1 GI:187098
VERSION M14180.1
KEYWORDS low density lipoprotein receptor-1.
SEGMENT 2 of 2
SOURCE Human white blood cell DNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS Hobbs, H.H., Brown, M.S., Goldstein, J.L. and Russell, D.W.
TITLE Deletion of exon encoding cysteine-rich repeat of low density lipoprotein receptor alters its binding specificity in a subject with familial hypercholesterolemia
J. Biol. Chem. 261 (28), 13114-13120 (1986)
JOURNAL MEDLINE 87008518
COMMENT Analysis of the LDL-receptor gene of a patient with familial hypercholesterolemia (FH) revealed the deletion of exon 5 resulting from a homologous recombination between repetitive Alu sequences of intron 4 and intron 5.

FEATURES
source
Location/Qualifiers
1..97
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="19p13.3"
join(M14178.1:1..97,1..97)
/gene="LDLR"
<1..>97
/gene="LDLR"
/note="LDLR intron E; G00-119-362"
42..72
/gene="LDLR"
/note="deletion target sequence"

BASE COUNT
19 a 34 c 25 g 19 t
ORIGIN Chromosome 19p13.2-p13.1.

Query Match
Best Local Similarity 100.0%; Score 31; DB 9; Length 97;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 22242 CCTGTATCCACACACTTTGGAGCCCAAG 22272
|||||
Db 77 CCTGTATCCACACACTTTGGAGCCCAAG 47

RESULT 13
HUMDLRDJ/c
LOCUS
DEFINITION Human familial hypercholesterolemia 626-a gene with a deletion of exon 5.
ACCESSION M14179
VERSION M14179.1 GI:187101
KEYWORDS low density lipoprotein receptor-1; very low density lipoprotein. Human fibroblast DNA, from patient FH-626.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS Hobbs, H.H., Brown, M.S., Goldstein, J.L. and Russell, D.W.
TITLE Deletion of exon encoding cysteine-rich repeat of low density lipoprotein receptor alters its binding specificity in a subject with familial hypercholesterolemia
J. Biol. Chem. 261, 13114-13120 (1986)
JOURNAL MEDLINE 87008518
COMMENT Analysis of the LDL-receptor gene of a patient with familial hypercholesterolemia (FH) revealed the deletion of exon 5 resulting from a homologous recombination between repetitive Alu sequences of intron 4 and intron 5.

FEATURES
source
Location/Qualifiers
1..97
/organism="Homo sapiens"
/db_xref="taxon:9606"

misc_feature
42..72
/note="deletion target sequence"
BASE COUNT
22 a 34 c 23 g 18 t
ORIGIN Chromosome 19p13.2-p13.1.

Query Match
Best Local Similarity 100.0%; Score 31; DB 9; Length 97;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 22242 CCTGTATCCACACACTTTGGAGCCCAAG 22272
|||||
Db 77 CCTGTATCCACACACTTTGGAGCCCAAG 47

RESULT 14
HUMALCE221/c
LOCUS
DEFINITION Human carcinoma cell-derived Alu RNA transcript, clone CE221.
ACCESSION M87896
VERSION M87896.1 GI:174874
KEYWORDS Alu repeat.
SOURCE Homo sapiens male embryo carcinoma cDNA to other RNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS Sinnett, D., Richer, C., Deragon, J.-M. and Labuda, D.
TITLE Alu RNA transcripts in human embryonal carcinoma cells. Model of post-transcriptional selection of master sequences
J. Mol. Biol. (1992) In press

FEATURES
source
Location/Qualifiers
1..103
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_line="NTera2D1"
/dev_stage="embryo"
/sex="male"
/tissue_type="carcinoma"

BASE COUNT
25 a 27 c 33 g 18 t
ORIGIN

Query Match 0.1%; Score 31; DB 9; Length 103;
 Best Local Similarity 100.0%; Pred. No. 0.00065;
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 17128 ATCTGCGCTCACTGCAACTCCGCTCCCG 17158
 ||||||||||||||||||||||||||||||||
 DB 85 ATCTGCGCTCACTGCAACTCCGCTCCCG 55

RESULT 15

HUMBRKFAA 80 bp DNA PRI 27-APR-1993
 LOCUS HUMBRKFAA 80 bp DNA PRI 27-APR-1993
 DEFINITION Human alpha-galactosidase breakpoint region.
 ACCESSION M36131
 VERSION M36131.1 GI:179541
 KEYWORDS Fabry disease; breakpoint junction; glycosphingolipid catabolism.
 SOURCE Human DNA.
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS Kornreich, R., Bishop, D.F. and Desnick, R.J.
 TITLE Alpha-galactosidase A gene rearrangements causing Fabry disease:
 identification of short direct repeats at breakpoints in an
 Alu-rich gene

JOURNAL J. Biol. Chem. 265, 9319-9326 (1990)
 MEDLINE 90264427

FEATURES
 source location/Qualifiers
 1..80
 /organism="Homo sapiens"
 /db_xref="taxon:9606"

BASE COUNT 19 a 28 c 20 g 13 t
 ORIGIN

Query Match 0.1%; Score 30; DB 9; Length 80;
 Best Local Similarity 100.0%; Pred. No. 0.0021;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 16511 GCCCACCCTTGCGCTCCCAAGTGTGGGAT 16540
 ||||||||||||||||||||||||||||
 DB 12 GCCCACCCTTGCGCTCCCAAGTGTGGGAT 41

Search completed: March 30, 2000, 23:41:00
 Job time: 159409 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 30, 2000, 17:38:54 ; Search time 612.33 Seconds
(without alignments)
12258.117 Million cell updates/sec

Title: US-08-852-495c-1_COPY_115000_145000
Perfect score: 30001
Sequence: 1 AATCTTCGAATTTGTATA.....AGAAATAACCTGATTTAA 30001

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 311585 seqs, 125096042 residues

Word size : 0

Total number of hits satisfying chosen parameters: 433264

Minimum DB seq length: 8
Maximum DB seq length: 105

Post-processing: Listing first 45 summaries

Database : N_Geneseq_36.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	28	0.1	40	1 Q55168	Sequence of primer
C 2	28	0.1	48	1 T04081	Trypsin inhibitory
C 3	28	0.1	51	1 T25038	Human gene signatu
C 4	28	0.1	57	1 Q95177	Simple tandem repe
C 5	28	0.1	60	1 T69401	PCR primer AL1. Nu
C 6	28	0.1	86	1 Q11760	Self-complementary
C 7	28	0.1	98	1 X00174	Porcine reproducti
C 8	27	0.1	40	1 V19044	Alu PCR primer 1.
C 9	27	0.1	61	1 T23178	Human gene signatu
C 10	27	0.1	80	1 V37197	Oligonucleotide se
C 11	27	0.1	88	1 T83744	Microsatellite ana
C 12	27	0.1	92	1 T88439	Padlock probe alpn
C 13	27	0.1	95	1 Q75099	Plasmid pOKSC18a c
C 14	27	0.1	100	1 V68827	DNA molecule encod
C 15	26	0.1	27	1 V71935	Anchored poly T RT
C 16	26	0.1	32	1 Q27389	Inter-Alu specific
C 17	26	0.1	35	1 Q27391	Inter-Alu specific
C 18	26	0.1	40	1 V19045	Alu PCR primer 2.
C 19	26	0.1	48	1 Q86183	Primer SINKball1700
C 20	26	0.1	48	1 T30807	Sindbis PCR primer
C 21	26	0.1	48	1 T35073	Sindbis-based, tum
C 22	26	0.1	48	1 V42384	Reverse PCR primer
C 23	26	0.1	48	1 V70704	Reverse PCR primer
C 24	26	0.1	69	1 V02148	Human secreted pro
C 25	26	0.1	69	1 T88081	3' portion of cDNA
C 26	26	0.1	79	1 V26706	Human novel secret
C 27	26	0.1	79	1 V32414	Homo sapiens clone
C 28	26	0.1	85	1 V05720	Nucleotide sequenc
C 29	26	0.1	92	1 T39467	Growth regulatory
C 30	26	0.1	92	1 V37485	Human growth regul
C 31	26	0.1	99	1 T91300	Human M97-2 secret
C 32	26	0.1	101	1 V00420	3' fragment of c10
C 33	25	0.1	27	1 T93831	Phosphodiester oli

C 34	25	0.1	29	1 V15487
C 35	25	0.1	37	1 T28941
C 36	25	0.1	37	1 V12343
C 37	25	0.1	41	1 V03013
C 38	25	0.1	46	1 T65781
C 39	25	0.1	47	1 T98523
C 40	25	0.1	48	1 Q93960
C 41	25	0.1	50	1 T39266
C 42	25	0.1	50	1 T92388
C 43	25	0.1	50	1 V26819
C 44	25	0.1	50	1 V37933
C 45	25	0.1	50	1 X02643

ALIGNMENTS

RESULT 1	
ID Q55168/C	055168 standard; DNA; 40 BP.
AC Q55168:	
DT 21-JUL-1994 (first entry)	
DE Sequence of primer for PCR amplification of HIV-LP Pt.1 isolate	
KW Human immunodeficiency virus; HIV-LP; PCR primer; ss.	
OS Synthetic.	
PN W09400562-A.	
PD 06-JAN-1994.	
PF 23-JUN-1993; U06162.	
PR 24-JUN-1992; US-903421.	
PA (CORR) CORNELL RES FOUND INC.	
PA (UANY) UNIV NEW YORK MT SINAI.	
PA SCHOOL MEDICINE.	
PI Gelman IH, Laurence JC;	
DR WPI; 94-02620/03.	
PT HIV-LP useful in vaccine formulations - is novel HIV virus	
PT distinct from HIV-1 and or HIV-2 viruses	
PS Example; page 7; 75pp; English.	
CC HIV-LP is a new variant of the HIV family. A cDNA first strand was	
CC synthesised from Pt. 1 pellet using MY RT. The product was	
CC converted into dsDNA and this cDNA was amplified by PCR using	
CC primers Q55167 and Q55168.	
SQ Sequence 40 BP; 3 A; 2 C; 3 G; 32 T;	

Query Match 0.1%; Score 28; DB 1; Length 40;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 29696 AAAAAAAAAAAAAAAAAAAGAAAT 29723
Db 34 AAAAAAAAAAAAAAAAAAAGAAAT 7

RESULT 2
ID T04081/C
ID T04081 standard; DNA; 48 BP.
AC T04081:
DT 16-MAY-1996 (first entry)
DE Trypsin inhibitory protein cDNA antisense primer-1.
KW Trypsin; inhibitor; human T98G cells; pancreatitis; shock; DIC;
KM multiple organ failure; disseminated intravascular coagulation; ss.
OS Synthetic.
PN J07242700-A.
PD 19-SEP-1995.
PF 04-MAR-1994; 059906.
PR 04-MAR-1994; JP-059906.
PA (MOCH) MOCHIDA PHARM CO LTD.
DR WPI; 95-355285/46.
PT Polypeptide having trypsin inhibitory activity - for the treatment
of pancreatitis, shock, multi-organ failure, etc.
PS Example 7; Page 15; 25pp; Japanese.
CC A polypeptide having trypsin inhibitory activity, mol.wt. 22-28 kD
or 17-23 kD (as determined by SDS-PAGE under reducing or non-

CC reducing conditions, respectively) and an N-terminal sequence as in
CC R79913 is claimed. The polypeptide was isolated from human T98G
CC cells and is useful for treating pancreatitis, shock, multiple
CC organ failure and disseminated intravascular coagulation. The
CC present sequence is that of an antisense primer used for cloning
CC T98G cDNA coding for the full-length polypeptide
SQ Sequence 48 Bp; 2 A; 1 C; 1 G; 44 T;

Query Match 0.1%; Score 28; DB 1; Length 48;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 29696 AAAAAAAAAAAAAAAAAAGAGAT 29723
DB 32 AAAAAAAAAAAAAAAAAAGAGAT 5

RESULT 3
ID T25038/c standard; cDNA to mRNA; 51 Bp.
AC T25038;
DT 11-NOV-1996 (first entry)
DE Human gene signature HUMG507164.
KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.
OS Homo sapiens.
PN WO9514772-A1.
PD 01-JUN-1995.
PR 11-NOV-1994; J01915.
PR 12-NOV-1993; JP-355504.
PA (HABS/) MATSUBARA K.
PI (ORUB/) OKUBO K.
PI Matsubara K, Okubo K;
DR WPI: 95-206931/27.
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues

CC Claim 1; Page 1755; 2245pp; Japanese.
CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in T19001-T26837 and which is able to hybridize to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridize with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.
SQ Sequence 51 Bp; 14 A; 16 C; 12 G; 9 T;

Query Match 0.1%; Score 28; DB 1; Length 51;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 16309 GCCCAGGCTGAGTGCAGTGCATGATC 16336
DB 28 GCCCAGGCTGAGTGCAGTGCATGATC 1

RESULT 4
ID O95177 standard; DNA; 57 Bp.
AC O95177;
DT 08-FEB-1996 (first entry)

DE Simple tandem repeat (STR) corresponding to the 2nd part of wglaz.
KW Simple tandem repeat; STR; wglaz; treatment; genetic; diagnosis;
KW characterisation; mapping; linkage studies; analysis; alleles;
KW second part; ss.
OS Synthetic.
PN WO9517522-A2.
PD 29-JUN-1995.
PR 21-DEC-1993; GB-026052.
PA (UYLE-) UNIV. LEICESTER.
PI Armour J, Jeffreys AJ;
DR WPI: 95-240682/31.
PT Identifying simple tandem repeat loci in DNA - by screening DNA
PT library to enrich for fragments contg. the repeats before cloning
PT and rescreening, also simple tandem repeats for treatment or
PT diagnosis

CC Claim 26; Page 17; 51pp; English.
CC O95177 is a simple tandem repeat (STR) corresponding to the 2nd part
CC of wglaz. The STR can be used for treatment and diagnosis in human
CC and veterinary medicine, partic. for genetic characterisation,
CC mapping, linkage studies and analysis/diagnosis of acquired disease
CC alleles.
SQ Sequence 57 Bp; 7 A; 11 C; 18 G; 21 T;

Query Match 0.1%; Score 28; DB 1; Length 57;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 16309 GCCCAGGCTGAGTGCAGTGCATGATC 16336
DB 30 GCCCAGGCTGAGTGCAGTGCATGATC 57

RESULT 5
ID T69401/c standard; DNA; 60 Bp.
AC T69401;
DT 26-AUG-1997 (first entry)
DE PCR primer ALI.
KW Pheromone receptor; vomeronasal sensory neuron; social behaviour;
KW maternal behaviour; reproductive behaviour; fertility;
KW hormone secretion; polymerase chain reaction; PCR; primer; ss.
OS Synthetic.
PN WO9714790-A1.
PD 24-APR-1997.
PR 18-OCT-1996; U16637.
PR 19-OCT-1995; US-005698.
PA (UYCO) UNIV COLUMBIA NEW YORK.
PI Axel R, Dulac C;
DR WPI: 97-245107/22.
PT Nucleic acid molecule encoding vertebrate pheromone receptor -
PT useful to identify modulators for control of reproductive and social
PT behaviour, fertility and hormone secretion
PS Disclosure: Page 30; 13pp; English.
CC main olfactory epithelium (MOE) and vomeronasal organ (VNO)
CC neurons contained an accurate representation of sequences in the
CC mRNA. Amplified cDNAs from single cells were used as probes,
CC leading to the isolation of a clone (T69545) coding for rat
CC pheromone receptor VN1 (W19103).
SQ Sequence 60 Bp; 11 A; 8 C; 8 G; 33 T;

Query Match 0.1%; Score 28; DB 1; Length 60;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 29696 AAAAAAAAAAAAAAAAAAGAGAT 29723
DB 60 AAAAAAAAAAAAAAAAAAGAGAT 33

```
RESULT 6
ID 011760/c
ID 011760 standard; DNA: 86 BP.
AC 011760:
DE 22-JUL-1991 (first entry)
DE Self-complementary, T7 promoter hairpin-forming sequence #2.
KW T7 bacteriophage; DNA-dependent RNA polymerase;
KW target sequence amplification; ss.
OS Synthetic.
FH Key
FH Location/Qualifiers
FT stem_loop 1..86
FT /*tag= a
FT /note= "number of T's in loop is 50"
EP-427073-A.
PD 15-MAY-1991.
PF 27-OCT-1990; 120650.
PR 09-NOV-1989; US-434372.
PR 23-AUG-1990; US-56991.
PA (MOLE-) MOLECULAR DIAGNOSTI.
PI Datta Gupta N;
PI WPI: 91-141638/20.
DR Nucleic acid probe for amplification and detection of target
PT sequence - capable of forming ligatable hairpin structured
PR promoter and transcribing target sequence, is sensitive and
PT useful in medical diagnosis
PS Claim 2; Page 11; 15pp; English.
CC The sequence is an example of a preferred T7 RNA polymerase
CC promoter for use in the invention. There can be from 2 to 50 T's in
CC the loop region. A probe sequence is ligated to the 3' end of the
CC promoter region. Upon hybridisation of the probe to a target sequence
CC and ligation of the hybridised target sequence to the 5' end of the
CC hairpin-forming sequence, the target sequence can be transcribed by
CC T7 RNA polymerase. This allows the target sequence to be amplified
CC (and detected) using a single oligonucleotide component (c.f. PCR).
CC See also 011759 and 011761-011764.
SQ Sequence 86 BP; 12 A; 6 C; 6 G; 62 T;

Query Match 0.1%; Score 28; DB 1; Length 86;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 29692 TATTAATAAAAAAAAAAAAAAAAAA 29719
DB 73 TATTAATAAAAAAAAAAAAAAAAAA 46

RESULT 7
ID X00174/c
ID X00174 standard; DNA: 98 BP.
AC X00174:
DE 23-MAR-1999 (first entry)
DE Porcine reproductive and respiratory syndrome virus PCR primer PRRSV-2.
KW Equine arteritis virus; EAV; vaccine; structural gene; PRRSV;
KW porcine reproductive and respiratory syndrome virus; recombinant virus;
KW PCR primer; ss.
OS Synthetic.
OS Porcine reproductive and respiratory syndrome virus.
PN WO9855626-A2.
PD 10-DEC-1998.
PE 05-JUN-1998; U12141.
PR 05-JUN-1997; US-048662.
PA (ORIG-) ORIGEN INC.
PI Consensus PM, Reilly JD, Spatz SJ;
PI WPI: 99-080829/07.
DR New recombinant porcine reproductive and respiratory syndrome virus
PT - containing nucleic acid encoding a polymerase from an RNA virus
PT and open reading frames 2-7 of the porcine virus, used particularly
PT in vaccines
PS Example 2; Page 21; 55pp; English.
CC The present invention describes a nucleic acid which encodes a
CC polymerase from an RNA virus, excluding porcine reproductive and
```

```
CC respiratory syndrome virus (PRRSV), and open reading frames (ORFs) 2-7
CC of PRRSV. The use of a polymerase gene from RNA viruses can provide for
CC production of less mutagenic recombinant viruses. The recombinant
CC viruses can be used in vaccines which have a reduced risk of loss or
CC reduction of efficacy. The vaccines are used particularly for
CC protecting swine against PRRSV. The high fidelity RNA polymerase gene
CC can be used as a marker that allows organisms vaccinated with such a
CC vaccine to be distinguished from organisms naturally infected with wild
CC type strains of virus or other vaccines. The present sequence represents
CC a PCR primer used in an example from the present invention.
SQ Sequence 98 BP; 11 A; 14 C; 8 G; 65 T;

Query Match 0.1%; Score 28; DB 1; Length 98;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 29696 AAAAAAAAAAAAAAAAAAAGNAT 29723
DB 32 AAAAAAAAAAAAAAAAAAAGNAT 5

RESULT 8
ID V19044
ID V19044 standard; DNA: 40 BP.
AC V19044:
DE 28-JUL-1998 (first entry)
DE Alu PCR primer 1.
KW PCR; primer; amplification; Alu repeat sequence; vector;
KW circular yeast artificial chromosome; YAC; ss.
OS Synthetic.
OS Saccharomyces sp.
PN WO9801573-A1.
PD 15-JAN-1998.
PE 09-JUL-1996; U11478.
PR 09-JUL-1996; WO-U11478.
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
PI Kourilina NY, Laktionov VL, Perkins EL, Resnick MA;
PI WPI: 98-110234/10.
DR Preparation of yeast artificial chromosomes - by in vivo
PT recombination using vector comprising yeast centromere, marker,
PT repeat telomere and nucleic acid for recombination
PS Example 1; Page 45; 117pp; English.
CC This is the nucleotide sequence for the PCR primer used in the
CC amplification of the Alu repeat sequence, which is used to
CC demonstrate the processes described in the invention. It involves
CC the creation and use of circular yeast artificial chromosome (YAC)
CC to selectively clone specific nucleic acids from a background of
CC mixed nucleic acids by introducing the vector(s) into E. coli cells.
CC They can be used to rapidly isolate human DNA where only a part of the
CC sequence of DNA is known. Using the methods large fragments of DNA can
CC be easily cloned and analysed.
SQ Sequence 40 BP; 7 A; 12 C; 13 G; 8 T;

Query Match 0.1%; Score 27; DB 1; Length 40;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 22242 CCTGTAATCCAGCACTTGGAGGCC 22268
DB 12 CCTGTAATCCAGCACTTGGAGGCC 38

RESULT 9
ID T23178
ID T23178 standard; cDNA to mRNA: 61 BP.
AC T23178:
DE 30-AUG-1996 (first entry)
DE Human gene signature HUGS04920.
KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.
```

OS Homo sapiens.
 PN W09514772-A1.
 PD 01-JUN-1995.
 PF 11-NOV-1994: J01916.
 PR 12-NOV-1993: JP-355504.
 PA (MARS/) MATSUBARA K.
 PI (OKUBO/) OKUBO K.
 PT Matsubara K, Okubo K;
 PT WPI: 95-206931/27.
 PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
 PT for diagnosis of abnormal cell function. By preparing cDNA that
 PT reflects relative abundance of corresp. mRNA in specific human
 PT tissues
 PS Claim 1; Page 1305; 2245BP; Japanese.
 CC A single-stranded DNA (or its complementary strand or the corresp.
 CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
 CC given in T19001-T26837 and which is able to hybridise to part of
 CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
 CC sequences were obtained from 3'-directed cDNA libraries prepared
 CC from various human tissues; synthesis of cDNA was initiated from the
 CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
 CC untranslated sequence is unique to a particular mRNA species, almost
 CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
 CC is constructed so as to reflect accurately the relative abundance of
 CC different mRNAs in the particular tissue from which it was derived.
 CC The appearance frequency of a given GS in a cDNA library can be
 CC determined (esp. using primers and probes derived from the GS
 CC sequences) as a means of diagnosing abnormal cell function or for
 CC recognising different cell types.
 SQ Sequence 61 BP; 23 A; 12 C; 14 G; 12 T;
 Query Match 0.1%; Score 27; DB 1; Length 61;
 Best Local Similarity 100.0%; Pred. No. 29;
 Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 18097 GATCAGCGACATGATTCAGCTGGG 18123
 DB 1 GATCAGCGACATGATTCAGCTGGG 27
 RESULT 10
 V37197/c
 ID V37197 standard; DNA: 80 BP.
 AC V37197.
 DT 14-SEP-1998 (first entry)
 DE Oligonucleotide sequence of the specification.
 KW Genotype; phenotype; molecular evolutionary engineering;
 KW functional biopolymer; virus; ss.
 OS Synthetic.
 PN W09816636-A1.
 PD 23-APR-1998.
 PF 17-OCT-1997: J03766.
 PR 17-OCT-1996: JP-274855.
 PA (MITU) MITSUBISHI CHEM CORP.
 PI Fusimi Y, Miyamoto E, Nemoto N, Yanagawa H;
 PT WPI: 96-261039/23.
 PT Virus containing nucleic acid and protein sections - for use in
 PT modification and creation of functional bio:polymers such as
 PT enzymes, antibodies and ribozyme(s).
 PS Example 1; Page 44; 68BP; Japanese.
 CC The present sequence is used in the course of the invention. The
 CC specification describes a molecule for bringing together genotype with
 CC phenotype (in vitro virus). The molecule contains a nucleic acid
 CC fragment having a base sequence corresponding to a genotype, covalently
 CC bonded to a protein fragment containing a protein participating in
 CC phenotype expression, the 3'-end of the nucleic acid part being bonded
 CC to the C-terminus of the protein part via a purmycin moiety. The nucleic
 CC acid fragment preferably consists of RNA corresponding to the gene (free
 CC from a termination codon), a spacer (such as two-stranded DNA), a peptide
 CC adapter (such as a DNA-tRNA hybrid containing an anticodon corresponding
 CC to the terminator codon of the gene) and a purmycin-containing cap
 CC capable of binding to an amino acid residue. Translation of the virus is

CC performed in a non-cellular (preferably ribosomal) system (e.g. E. coli
 CC ribosome), the protein synthesised by the translation attaching to the
 CC purmycin cap, resulting in the complete in vitro virus structure. The
 CC method is used in molecular evolutionary engineering to optimise function
 CC of a functional biopolymer such as an enzyme, antibody or ribozyme, or
 CC to generate new functionality. The virus can be used for the functional
 CC optimisation of nucleic acid or protein sequences.
 SQ Sequence 80 BP; 7 A; 10 C; 1 G; 62 T;
 Query Match 0.1%; Score 27; DB 1; Length 80;
 Best Local Similarity 100.0%; Pred. No. 27;
 Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 29693 ATTMAAAAAAAAAAAAAAAAAAAAAA 29719
 DB 66 ATTMAAAAAAAAAAAAAAAAAAAAAA 40
 RESULT 11
 V39744
 ID V39744 standard; cDNA: 88 BP.
 AC V39744.
 DT 28-SEP-1998 (first entry)
 DE Microsatellite analysis antisense strand SEQ ID NO:262.
 KW Mass spectrometry; diagnosis; detection; biological sample; infection;
 KW genetic disease; chromosomal abnormality; identification; heredity;
 KW pathogenic organism; telomerase activity; oncogene mutation;
 KW cancer-specific sequence; primer; ss.
 OS Synthetic.
 PN W09820166-A2.
 PD 14-MAY-1998.
 PF 06-NOV-1997: U20444.
 PR 08-OCT-1997: US-947801.
 PR 06-NOV-1996: US-744481.
 PR 06-NOV-1996: US-744590.
 PR 06-NOV-1996: US-746036.
 PR 06-NOV-1996: US-746035.
 PR 23-JAN-1997: US-786988.
 PR 23-JAN-1997: US-787639.
 PR 19-SEP-1997: US-933792.
 PA (SEOU-) SEQUENOM INC.
 PI Braun A, Damhofer-Demar B, Fu D, Higgins GS, Jurinke C,
 PI Koster H, Little DP, Lough DW, Siebert CW, Tang X,
 PI Van Den Boom D, Xiang G;
 PT WPI: 98-286975/23
 DT Sequencing nucleic acid by mass spectrometric analysis - for
 DT detecting nucleic acids, telomerase activity, oncogene mutations, or
 DT cancer-specific sequences for diagnosis of disease
 PS Example 11; Page 318; 478BP; English.
 CC A process has been developed for determining the sequence of a target
 CC nucleic acid. The process comprises: (i) generating at least two
 CC fragments (F) from the target nucleic acid; and (ii) analysing F by
 CC mass spectrometry (MS). The sequences in V39483 to V39592 are
 CC specifically claimed primers for use in the mass spectrometric analysis
 CC of the above process. The process is used to detect genetic diseases
 CC (e.g. haemophilia, thalassemia, Duchenne muscular dystrophy, Alzheimer's
 CC disease, cystic fibrosis and many others) or chromosomal abnormalities
 CC (or predisposition); infections and cancers; also for establishing
 CC neuroblastoma, detecting telomerase, determining family relationships
 CC and HLA compatibility, and in genetic fingerprinting. Compared with
 CC known methods using MS, this process requires fewer specific reagents
 CC and is better suited to automation. Extended primers are shorter
 CC primer annealing is more efficient and the process allows detection of
 CC many sequences simultaneously. The present sequence represents an
 CC oligonucleotide used in an example from the present invention.
 SQ Sequence 88 BP; 36 A; 5 C; 7 G; 20 T;
 Query Match 0.1%; Score 27; DB 1; Length 88;
 Best Local Similarity 100.0%; Pred. No. 26;
 Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
OY 23274 AAAAATAAATAAATAAATAA 23300
Db 15 AAAAATAAATAAATAAATAA 41

RESULT 12
T88439/c
ID T88439 standard; DNA: 92 BP.
AC T88439;
DE 12-MAY-1998 (first entry)
DE Padlock probe alphaFV from WO9741254.
KM Cleavable padlock probe; target nucleic acid; detection;
KM circularised structure; hybridisation; Factor V; ss.
OS Synthetic.
FH Key Location/Qualifiers
FT modified_base 67
FT /tag= a
FT /note= "attached to an amino group"
PD WO9741254-A1.
PD 06-NOV-1997.
PR 30-APR-1997; SE0737.
PR 30-APR-1996; SE-001676.
PA (LAND/) LANDEGREN U.
PI Landegren U;
PI WPI: 97-549751/50.
DR Detecting target nucleic acid sequence - using probe capable of
PT forming circularised structure, which can be cleaved to form
PT detectable function
PS Example 2; Fig 7; 20pp; English.
CC The present invention describes a method for detecting a target nucleic
CC acid sequence (I) in a sample. The method comprises: (a) hybridising (I)
CC to the probe ends of a probe having 2 free nucleic acid parts, which
CC are at least partially complementary to and capable of hybridising to
CC at least 2 neighbouring regions of (I); (b) covalently connecting the
CC ends of the hybridised probe with each other to form a circularised
CC structure; (c) washing under denaturing conditions, characterised in
CC that the probe is provided with a cleavable or dissociable detectable
CC function; (d) cleaving or dissociating the detectable function; (e)
CC separating probes with connected ends from probes with non-connected
CC ends by washing under denaturing conditions; and (f) detecting the
CC presence and, if desired, location of the remaining probe as indicative
CC of the presence of (I). The present sequence represents a padlock probe,
CC alpha FV, used in an example of the present invention. The method
CC can be used to detect, quantify and distinguish between sequence
CC variants with regard to 1 or several (I) in a sample. It can be used to
CC distinguish between normal and mutated sequence variants associated with
CC disease, for genetic linkage analysis of biallelic markers and to
CC quantify gene expression in a tissue sample. The circularisable probe is
CC designed so that it reports the presence of (I) by allowing a detectable
CC moiety to remain bound only if the probe has been cyclised in a target
CC dependent linking reaction. It also provides for high specificity and
CC background reduction.
CC Sequence 92 BP; 9 A; 9 C; 8 G; 66 T;
SQ
```

```
Query Match 0.1%; Score 27; DB 1; Length 92;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 23696 AAAAAAAAAAAAAAAAAAGAA 29722
Db 69 AAAAAAAAAAAAAAAAAAGAA 43
```

```
RESULT 13
ID 075099/c
AC 075099;
DE 11-AUG-1995 (first entry)
DE Plasmid pOKS18a confg. portion of psychosis protecting protein gene.
KM Psychosis protecting protein; PCR; primer; ss.
OS Homo sapiens.
```

```
FH Key Location/Qualifiers
FT misc_difference 1..28
FT /tag= a
FT /note= "PCR primer used to generate cDNA library"
FT misc_difference 74..95
FT /tag= b
FT /note= "see above"
PD WO9426107-A.
PD 24-NOV-1994.
PD 13-MAY-1994; 005445.
PR 13-MAY-1993; US-060560.
PA (UYNV) UNIV NEW YORK STATE.
PI Basham DA, Friedhoff AJ, Miller JC;
PI WPI: 95-006234/01.
DR New nucleic acids encoding psychosis protecting peptide and
PT antibodies - for the treatment, diagnosis and research of
PT psychotic disorders, such as schizophrenia
PS Claim 1; Page 52; 87 pp; English.
CC A subtracted cDNA library was generated from the RNA of lymphocytes
CC obtd. from monozygotic twins discordant for schizophrenia. The two
CC oligos used in library construction were T7 promoter and SP6
CC promoter (see 075091, 075092). A cDNA library was made for each
CC twin and subtractive hybridisation was achieved. Both libraries
CC were used as driver and substrate in two separate subtraction
CC assays. Clone pOKS18a was isolated and sequenced.
SQ Sequence 95 BP; 12 A; 14 C; 30 G; 39 T;
```

```
Query Match 0.1%; Score 27; DB 1; Length 95;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 24134 GAAGTCGATCTCAAAAAAAAAA 24160
Db 41 GAAGTCGATCTCAAAAAAAAAA 15
```

```
RESULT 14
ID V68827/c
AC V68827;
DE 22-JAN-1999 (first entry)
DE DNA molecule encoding a breast tumour specific polypeptide #19.
KM Human; breast cancer; breast tumour tissue; diagnosis; treatment;
KM vaccine; epitope; endogenous; retroviral element; ss.
OS Homo sapiens.
PD WO9845328-A2.
PD 15-OCT-1998.
PR 09-APR-1998; 006939.
PR 11-DEC-1997; US-991789.
PR 09-APR-1997; US-838762.
PA (CORI-) CORIXA CORP.
PI Fridakis TN, Reed SG, Smith JM;
PI WPI: 98-557473/47.
DR New DNA sequences isolated from endogenous human retroviral element
PT - and related vectors, transformed cells, proteins and antibodies,
PT useful for diagnosis, treatment and prevention of breast cancer
PS Claim 1; Page 47; 173pp; English.
CC V68800 to V68998 represent nucleotide sequences which encode human
CC breast tumour specific polypeptides. Detection or measurement of
CC human breast tumour specific polypeptides and nucleotide sequences,
CC or the corresponding RNA in a sample, is used for diagnosis and
CC monitoring of breast cancer. Human breast tumour specific polypeptides
CC and nucleotide sequences, and the vectors containing the DNAs, are also
CC useful in vaccines for inhibiting development (for prevention or
CC therapy) of breast cancer. The polypeptides may also be used to
CC raise monoclonal antibodies, used as immunoassay reagents.
SQ Sequence 100 BP; 19 A; 10 C; 8 G; 63 T;
```

```
Query Match 0.1%; Score 27; DB 1; Length 100;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Mon Apr 3 08:24:03 2000

us-08-852-495c-1_copy_115000_145000.rng

OY 29696 AAAAAAAAAAAAAAAAAAGAA 29722
DB 74 AAAAAAAAAAAAAAAAAAGAA 48

RESULT 15

V71935/C
ID V71935 standard; DNA; 27 BP.
AC V71935; 18-FEB-1999 (first entry)
DE Anchored poly (RT-PCR primer).
KW Normalised; cDNA library; mRNA cloning; reverse transcription;
KW Immobilised; screening; hybridisation; nucleic acid amplification;
KW expression pattern; drug development; PCR primer; RT-PCR; ss.
OS Synthetic.
PN W09851789-A2.
PD 19-NOV-1998.
PE 13-MAY-1998; DK0186.
PR 27-MAR-1998; DK-000432.
PR 13-MAY-1997; DK-000547.
PR 19-MAY-1997; US-871030.
PA (DISP-) DISPLAY SYSTEMS BIOTECH APS.
PI Marchoe PR.
DR WPI: 99-009772/01.
PR Preparation of normalised, subdivided cDNA libraries from mRNA - by
PR reverse transcription and amplification, used to screen for new
PR genes and interacting proteins. Potential drugs, and for diagnosis
PR Example 1; Page 29; 71pp; English.
CC The invention relates to preparation of a normalised, subdivided library
CC of amplified cDNA from the coding regions of mRNA in a sample. The method
CC involves reverse transcription with at least one cDNA primer of formula
CC of 5'-d(Gn2-Vn3-Nn4) to form first strand cDNA where Gn1 - any sequence
CC are 100 nucleotides; dr - deoxythymidyl, n2 is at least 1; n3 and n4
CC cDNA 5'-OH, or n3 is 1 and n4 is at least 1; followed by second strand
CC primer synthesis using the first strand as template and a second cDNA
CC primer of a similar formula, in the presence of DNA polymerase I (or its
CC amplification primers. Comparison of cDNA in the prepared library with a
CC database of computer-generated list of molecular weights of restricted
CC DNA fragments of known sequence) is used to determine presence of an
CC expressed protein in a cell, also to detect changes in such expression.
CC (particularly for diagnosis of disease). Surfaces (chip) having amplified
CC cDNA stably immobilised on it, obtained by a similar method, are used
CC to screen for genes of a particular family, by hybridisation with nucleic
CC acid from the family (to identify new genes) and to detect differences in
CC expression patterns between cells. The polypeptides expressed by the
CC libraries can be used for drug development. Sequences V71935 to V71946
CC represent primers used to exemplify the method of the invention.
SO Sequence 27 BP; 2 A; 0 C; 0 G; 25 T;

Query Match 0.18; Score 26; DB 1; Length 27;
Best Local Similarity 100.0%; Pred. No. 67;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 29694 TTTAAAAAAAAAAAAAAAAAAAAA 29719
DB 27 TTTAAAAAAAAAAAAAAAAAAAAA 2

Search completed: March 31, 2000, 00:01:30
Job time: 159088 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 30, 2000, 17:28:35 ; Search time 345.24 Seconds

(without alignments)
10403.818 Million cell updates/sec

Title: US-08-852-495C-1_COPY_115000_145000

Perfect score: 30001
Sequence: 1 AATCTTCCAAATTTTGATA.....AGAAATACCTATTAA 30001

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 214294 seqs, 59861574 residues

Word size : 0

Total number of hits satisfying chosen parameters: 349160

Minimum DB seq length: 8

Maximum DB seq length: 105

Post-processing: Listing first 45 summaries

Database :
1: /cgn2_6/ptodata/1/lna/5A.COMB.seq:*
2: /cgn2_6/ptodata/1/lna/5B.COMB.seq:*
3: /cgn2_6/ptodata/1/lna/5C.COMB.seq:*
4: /cgn2_6/ptodata/1/lna/5D.COMB.seq:*
5: /cgn2_6/ptodata/1/lna/6.COMB.seq:*
6: /cgn2_6/ptodata/1/lna/6_PCT059.COMB.seq:*
7: /cgn2_6/ptodata/1/lna/Backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	38	0.1	60	3	US-08-454-557C-57	Sequence 57, Appl
2	38	0.1	60	4	US-08-340-426D-57	Sequence 57, Appl
3	38	0.1	60	4	US-08-450-673C-57	Sequence 57, Appl
4	38	0.1	60	6	PCR-US95-17111A-57	Sequence 57, Appl
5	38	0.1	76	3	US-08-454-557C-69	Sequence 69, Appl
6	38	0.1	76	4	US-08-340-426D-69	Sequence 69, Appl
7	38	0.1	76	4	US-08-450-673C-69	Sequence 69, Appl
8	38	0.1	76	6	PCR-US95-17111A-69	Sequence 69, Appl
9	28	0.1	50	1	US-08-233-609-5	Sequence 5, Appl
10	28	0.1	50	1	US-08-444-083-5	Sequence 5, Appl
11	28	0.1	50	1	US-08-286-304-5	Sequence 5, Appl
12	28	0.1	50	1	US-08-442-745-5	Sequence 5, Appl
13	28	0.1	50	1	US-08-443-129-5	Sequence 5, Appl
14	28	0.1	50	1	US-08-443-952-5	Sequence 5, Appl
15	28	0.1	50	2	US-08-443-130-5	Sequence 5, Appl
16	28	0.1	50	6	PCR-US95-04467-5	Sequence 5, Appl
17	28	0.1	57	3	US-08-332-766A-5	Sequence 5, Appl
18	26	0.1	35	1	US-08-253-889-10	Sequence 10, Appl
19	26	0.1	42	1	US-07-875-167-2	Sequence 2, Appl
20	26	0.1	42	1	US-08-287-164-2	Sequence 2, Appl
21	26	0.1	48	2	US-08-741-881-21	Sequence 21, Appl
22	26	0.1	48	2	US-08-739-158-21	Sequence 21, Appl
23	26	0.1	48	3	US-08-739-167-21	Sequence 21, Appl
24	26	0.1	92	1	US-08-120-827-94	Sequence 94, Appl
25	26	0.1	92	2	US-08-478-675-94	Sequence 94, Appl
26	25	0.1	26	2	US-08-621-914A-1	Sequence 1, Appl

27	25	0.1	33	7	5478746-1	Patent No. 5478746
28	25	0.1	41	1	US-08-113-646A-39	Sequence 39, Appl
29	25	0.1	44	1	US-08-113-646A-40	Sequence 40, Appl
30	25	0.1	46	1	US-08-222-177A-349	Sequence 349, App
31	25	0.1	47	4	US-08-778-494B-114	Sequence 114, App
32	25	0.1	48	2	US-08-380-438-6	Sequence 6, Appl
33	25	0.1	50	1	US-08-381-572-20	Sequence 20, Appl
34	25	0.1	50	2	US-08-593-820-20	Sequence 20, Appl
35	25	0.1	50	4	US-08-663-823B-72	Sequence 72, Appl
36	25	0.1	51	4	US-08-582-562A-8	Sequence 8, Appl
37	25	0.1	51	4	US-08-778-494B-8	Sequence 8, Appl
38	25	0.1	51	4	US-08-859-998-1373	Sequence 1373, Ap
39	25	0.1	52	4	US-08-778-494B-111	Sequence 111, Appl
40	25	0.1	54	4	US-08-771-624B-24	Sequence 24, Appl
41	25	0.1	55	1	US-08-113-646A-41	Sequence 41, Appl
42	25	0.1	55	4	US-08-522-269B-18	Sequence 18, Appl
43	25	0.1	55	4	US-08-582-562A-16	Sequence 16, Appl
44	25	0.1	55	4	US-08-778-494B-16	Sequence 16, Appl
45	25	0.1	55	4	US-08-778-494B-67	Sequence 67, Appl

ALIGNMENTS

RESULT 1
US-08-454-557C-57
Sequence 57, Application US/08454557C
Patent No. 5630670
GENERAL INFORMATION:
APPLICANT: de la Monte, Suzanne
TITLE OF INVENTION: Neutral Thread Protein Gene Expression and Detection
TITLE OF INVENTION: of Alzheimer's Disease
NUMBER OF SEQUENCES: 121
CORRESPONDENCE ADDRESS:
ADDRESS: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/454,557C
FILING DATE: 30-MAY-1995
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609,3840003
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 57:
SEQUENCE CHARACTERISTICS:
LENGTH: 60 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
US-08-454-557C-57
Query Match 0.1%; Score 38; DB 3; Length 60;
Best Local Similarity 100.0%; Pred. No. 0.00064;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 17232 TTTTGTATTTTGTAGACAGGCTTCCACCGTCT 17269
DB 16 TTTTGTATTTTGTAGACAGGCTTCCACCGTCT 53

Mon Apr 3 08:24:04 2000

us-08-052-495c-1_copy_115000_145000.rml

Page 2

RESULT 2
US-08-340-426D-57
Sequence 57, Application US/08340426D
Patent No. 5948634
GENERAL INFORMATION:
APPLICANT: de la Monte, Suzanne
TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection
NUMBER OF SEQUENCES: 121
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/340,426D
FILING DATE: 14-NOV-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609.3840002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 57:
SEQUENCE CHARACTERISTICS:
LENGTH: 60 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
US-08-340-426D-57
Query Match 0.1%; Score 38; DB 4; Length 60;
Best Local Similarity 100.0%; Pred. No. 0.00064;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 17232 TTTTGTATTTTACTAGACAGGTTTCACCGTGT 17269
DB 16 TTTTGTATTTTACTAGACAGGTTTCACCGTGT 53
RESULT 3
US-08-450-673C-57
Sequence 57, Application US/08450673C
Patent No. 5948888
GENERAL INFORMATION:
APPLICANT: de la Monte, Suzanne
TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection
NUMBER OF SEQUENCES: 121
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/450,673C
FILING DATE: 30-MAY-1995
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609.3840004
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 57:
SEQUENCE CHARACTERISTICS:
LENGTH: 60 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
US-08-450-673C-57
Query Match 0.1%; Score 38; DB 4; Length 60;
Best Local Similarity 100.0%; Pred. No. 0.00064;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 17232 TTTTGTATTTTACTAGACAGGTTTCACCGTGT 17269
DB 16 TTTTGTATTTTACTAGACAGGTTTCACCGTGT 53
RESULT 4
PCT-US95-17111A-57
Sequence 57, Application PCT/US951711A
GENERAL INFORMATION:
APPLICANT: de la Monte, Suzanne
TITLE OF INVENTION: Neural Thread Protein Gene Expression and
NUMBER OF SEQUENCES: 121
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/17111A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/340,426
FILING DATE: 14-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609.3840002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 57:
SEQUENCE CHARACTERISTICS:
LENGTH: 60 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
PCT-US95-17111A-57

Query Match 0.1%; Score 38; DB 6; Length 60;

Best Local Similarity 100.0%; Pred. No. 0.00064;

Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17232 TTTTGTATTTTAGTAGACAGAGGTTTCACCGTGT 17269

DB 16 TTTTGTATTTTAGTAGACAGAGGTTTCACCGTGT 53

RESULT 5

US-08-454-557C-69

Sequence 69, Application US/08454557C

Patent No. 5830670

GENERAL INFORMATION:

APPLICANT: de la Monte, Suzanne

APPLICANT: Wands, Jack R.

TITLE OF INVENTION: Neutral Thread Protein Gene Expression and Detection

TITLE OF INVENTION: of Alzheimer's Disease

NUMBER OF SEQUENCES: 121

CORRESPONDENCE ADDRESS:

ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.

STREET: 1100 New York Avenue, Suite 600

CITY: Washington

STATE: D.C.

COUNTRY: U.S.A.

ZIP: 20005-3934

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/454,557C

FILING DATE: 30-MAY-1995

CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:

NAME: Ludwig, Steven R.

REGISTRATION NUMBER: 36,203

REFERENCE/DOCKET NUMBER: 0609.3840003

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 371-2600

TELEFAX: (202) 371-2540

INFORMATION FOR SEQ. ID NO: 69:

SEQUENCE CHARACTERISTICS:

LENGTH: 76 base pairs

TYPE: nucleic acid

STRANDEDNESS: both

TOPOLOGY: both

US-08-454-557C-69

Query Match 0.1%; Score 38; DB 3; Length 76;

Best Local Similarity 100.0%; Pred. No. 0.00061;

Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17232 TTTTGTATTTTAGTAGACAGAGGTTTCACCGTGT 17269

DB 16 TTTTGTATTTTAGTAGACAGAGGTTTCACCGTGT 53

RESULT 6

US-08-340-426D-69

Sequence 69, Application US/08340426D

Patent No. 5948634

GENERAL INFORMATION:

APPLICANT: de la Monte, Suzanne

APPLICANT: Wands, Jack R.

TITLE OF INVENTION: Neutral Thread Protein Gene Expression and Detection

TITLE OF INVENTION: of Alzheimer's Disease

NUMBER OF SEQUENCES: 121

CORRESPONDENCE ADDRESS:

ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.

STREET: 1100 New York Avenue, Suite 600

CITY: Washington

STATE: D.C.

COUNTRY: U.S.A.

ZIP: 20005-3934

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/340,426D

FILING DATE: 14-NOV-1994

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Ludwig, Steven R.

REGISTRATION NUMBER: 36,203

REFERENCE/DOCKET NUMBER: 0609.3840002

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 371-2600

TELEFAX: (202) 371-2540

INFORMATION FOR SEQ. ID NO: 69:

SEQUENCE CHARACTERISTICS:

LENGTH: 76 base pairs

TYPE: nucleic acid

STRANDEDNESS: both

TOPOLOGY: both

US-08-340-426D-69

Query Match 0.1%; Score 38; DB 4; Length 76;

Best Local Similarity 100.0%; Pred. No. 0.00061;

Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17232 TTTTGTATTTTAGTAGACAGAGGTTTCACCGTGT 17269

DB 16 TTTTGTATTTTAGTAGACAGAGGTTTCACCGTGT 53

RESULT 7

US-08-450-673C-69

Sequence 69, Application US/08450673C

Patent No. 5948888

GENERAL INFORMATION:

APPLICANT: de la Monte, Suzanne

APPLICANT: Wands, Jack R.

TITLE OF INVENTION: Neutral Thread Protein Gene Expression and Detection

TITLE OF INVENTION: of Alzheimer's Disease

NUMBER OF SEQUENCES: 121

CORRESPONDENCE ADDRESS:

ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.

STREET: 1100 New York Avenue, Suite 600

CITY: Washington

STATE: D.C.

COUNTRY: U.S.A.

ZIP: 20005-3934

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/450,673C

FILING DATE: 30-MAY-1995

CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:

NAME: Ludwig, Steven R.

REGISTRATION NUMBER: 36,203

REFERENCE/DOCKET NUMBER: 0609.3840004

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 371-2600

TELEFAX: (202) 371-2540

Mon Apr 3 08:24:04 2000

us-08-852-495c-1_copy_115000_145000.rn1

Page 4

INFORMATION FOR SEQ ID NO: 69:
SEQUENCE CHARACTERISTICS:
LENGTH: 76 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
US-08-450-673C-69

Query Match
Best Local Similarity 100.0%; Score 38; DB 4; Length 76;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 17232 TTTTGTATTTTGTAGAGACAGGTTTCACCGTGT 17269
DB 16 TTTTGTATTTTGTAGAGACAGGTTTCACCGTGT 53

RESULT 8
PCT-US95-17111A-69
Sequence 69, Application PC/TUS9517111A
GENERAL INFORMATION:
APPLICANT: de la Monte, Suzanne
TITLE OF INVENTION: Neural Thread Protein Gene Expression and
NUMBER OF SEQUENCES: 121
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/17111A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/340,426
FILING DATE: 14-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609,3840002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 69:
SEQUENCE CHARACTERISTICS:
LENGTH: 76 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
PCT-US95-17111A-69

Query Match
Best Local Similarity 100.0%; Score 38; DB 6; Length 76;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 17232 TTTTGTATTTTGTAGAGACAGGTTTCACCGTGT 17269
DB 16 TTTTGTATTTTGTAGAGACAGGTTTCACCGTGT 53

RESULT 9
US-08-233-609-5/C

Sequence 5, Application US/08233609
Patent No. 5534615
GENERAL INFORMATION:
APPLICANT: Baker, Joffre
APPLICANT: Chien, Kenneth
APPLICANT: King, Kathleen
APPLICANT: Pennica, Diane
APPLICANT: Wood, William
TITLE OF INVENTION: Cardiac Hypertrophy Factor and Uses Therefor
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: patln (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/233,609
FILING DATE: 25-APR-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Hasak, Janet E.
REGISTRATION NUMBER: 28,616
REFERENCE/DOCKET NUMBER: 894
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-1896
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 50 bases
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-233-609-5

Query Match
Best Local Similarity 100.0%; Score 28; DB 1; Length 50;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 29696 AAAAAAAAAAAAAAAAAAAAAAAGAAAT 29723
DB 44 AAAAAAAAAAAAAAAAAAAAAAAGAAAT 17

RESULT 10
US-08-444-083-5/C
Sequence 5, Application US/08444083
Patent No. 5571675
GENERAL INFORMATION:
APPLICANT: Baker, Joffre
APPLICANT: Chien, Kenneth
APPLICANT: King, Kathleen
APPLICANT: Pennica, Diane
APPLICANT: Wood, William
TITLE OF INVENTION: Cardiac Hypertrophy Factor and Uses Therefor
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA

ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/444,083
FILING DATE: 17-MAY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/233609
FILING DATE: 25-APR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/286304
FILING DATE: 05-AUG-1994
ATTORNEY/AGENT INFORMATION:
NAME: Torchia, Timothy E.
REGISTRATION NUMBER: 36,700
REFERENCE/DOCKET NUMBER: 894P1D5
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-8674
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 50 bases
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-444-083-5

Query Match 0.1%; Score 28; DB 1; Length 50;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 29696 AAAAAAAAAAAAAAAAAAGAAAT 29723
DB 44 AAAAAAAAAAAAAAAAAAGAAAT 17

RESULT 11
US-08-286-304-5/C
Sequence 5, Application US/08286304
Patent No. 5571893
GENERAL INFORMATION:
APPLICANT: Baker, Joffre
APPLICANT: Chien, Kenneth
APPLICANT: King, Kathleen
APPLICANT: Pennica, Diane
APPLICANT: Wood, William
TITLE OF INVENTION: Cardiac Hypertrophy Factor and Uses Therefor
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/286,304
FILING DATE: 05-AUG-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/233609
FILING DATE: 25-APR-1994

ATTORNEY/AGENT INFORMATION:
NAME: Hasak, Janet E.
REGISTRATION NUMBER: 28,616
REFERENCE/DOCKET NUMBER: 894P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-1896
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 50 bases
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-286-304-5

Query Match 0.1%; Score 28; DB 1; Length 50;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 29696 AAAAAAAAAAAAAAAAAAGAAAT 29723
DB 44 AAAAAAAAAAAAAAAAAAGAAAT 17

RESULT 12
US-08-442-745-5/C
Sequence 5, Application US/08442745
Patent No. 5624806
GENERAL INFORMATION:
APPLICANT: Baker, Joffre
APPLICANT: Chien, Kenneth
APPLICANT: King, Kathleen
APPLICANT: Pennica, Diane
APPLICANT: Wood, William
TITLE OF INVENTION: Cardiac Hypertrophy Factor and Uses Therefor
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/442,745
FILING DATE: 17-may-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/233609
FILING DATE: 25-APR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/286304
FILING DATE: 05-AUG-1994
ATTORNEY/AGENT INFORMATION:
NAME: Torchia, Timothy E.
REGISTRATION NUMBER: 36,700
REFERENCE/DOCKET NUMBER: 894P1D1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-8674
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 50 bases
TYPE: nucleic acid
STRANDEDNESS: single

Mon Apr 3 08:24:04 2000

us-08-852-495c-1_copy-115000_145000.rml

TOPOLOGY: linear

Query Match 0.1% Score 28; DB 1, Length 50;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 29696 AAAAAAAAAAAAAAAAAAGAAAT 29723
DB 44 AAAAAAAAAAAAAAAAAAGAAAT 17

US-08-443-129-5/c
Sequence 5, Application US/08443129

Patent No. 5627073

GENERAL INFORMATION:

APPLICANT: Baker, Joffre

APPLICANT: Chien, Kenneth

APPLICANT: King, Kathleen

APPLICANT: Pennica, Diane

APPLICANT: Wood, William

TITLE OF INVENTION: Cardiac Hypertrophy Factor and Uses Therefor

NUMBER OF SEQUENCES: 8

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Genentech, Inc.

STREET: 460 Point San Bruno Blvd

CITY: South San Francisco

STATE: California

COUNTRY: USA

ZIP: 94080

COMPUTER READABLE FORM:

MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patin (Genentech)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/443.129

FILING DATE: 17-MAR-1995

CLASSIFICATION: 1.35

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/233609

FILING DATE: 25-APR-1994

PRIOR APPLICATION DATA: 08/286

FILING DATE: 05-AUG-1994

PRIOR APPLICATION DATA: 304

FILING DATE: (null)

ATTORNEY/AGENT INFORMATION:

NAME: Torchia, Timothy B.

REGISTRATION NUMBER: 36,700

REFERENCE/DOCKET NUMBER: 894P1D2

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415/225-8674

TELEFAX: 415/952-9881

TELEX: 910/371-7168

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 50 bases

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

US-08-443-129-5

Query Match 0.1% Score 28; DB 1, Length 50;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 29696 AAAAAAAAAAAAAAAAAAGAAAT 29723
DB 44 AAAAAAAAAAAAAAAAAAGAAAT 17

US-08-443-952-5/c

Sequence 5, Application US/08443952
Patent No. 5679545

GENERAL INFORMATION:

APPLICANT: Baker, Joffre

APPLICANT: Chien, Kenneth

APPLICANT: King, Kathleen

APPLICANT: Pennica, Diane

APPLICANT: Wood, William

TITLE OF INVENTION: Cardiac Hypertrophy Factor and Uses Therefor

NUMBER OF SEQUENCES: 8

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Genentech, Inc.

STREET: 460 Point San Bruno Blvd

CITY: South San Francisco

STATE: California

COUNTRY: USA

ZIP: 94080

COMPUTER READABLE FORM:

MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patin (Genentech)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/443.952

FILING DATE: 17-MAR-1995

CLASSIFICATION: 4.35

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/233609

FILING DATE: 25-APR-1994

PRIOR APPLICATION DATA: 08/286304

FILING DATE: 05-AUG-1994

PRIOR APPLICATION DATA: 304

FILING DATE: (null)

ATTORNEY/AGENT INFORMATION:

NAME: Torchia, Timothy B.

REGISTRATION NUMBER: 36,700

REFERENCE/DOCKET NUMBER: 894P1D4

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415/225-8674

TELEFAX: 415/952-9881

TELEX: 910/371-7168

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 50 bases

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

US-08-443-952-5

Query Match 0.1% Score 28; DB 1, Length 50;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 29696 AAAAAAAAAAAAAAAAAAGAAAT 29723
DB 44 AAAAAAAAAAAAAAAAAAGAAAT 17

```

: TITLE OF INVENTION: Cardiac Hypertrophy Factor and Uses
: TITLE OF INVENTION: Therefor
: NUMBER OF SEQUENCES: 8
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Genentech, Inc.
: STREET: 460 Point San Bruno Blvd
: CITY: South San Francisco
: STATE: California
: COUNTRY: USA
: ZIP: 94080
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: 5.25 inch, 360 KB floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: patln (Genentech)
:
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/443,130
: FILING DATE: 17-MAY-1995
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/233609
: FILING DATE: 25-APR-1994
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/286304
: FILING DATE: 05-AUG-1994
: ATTORNEY/AGENT INFORMATION:
: NAME: Torchia, Timothy E.
: REGISTRATION NUMBER: 36,700
: REFERENCE/DOCKET NUMBER: 894PID3
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 415/225-8674
: TELEFAX: 415/952-9881
: TELEX: 910/371-7168
:
: INFORMATION FOR SEQ ID NO: 5:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 50 bases
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
:
: US-08-443-130-5

```

```

Query Match          0.1%; Score 28; D3 2; Length 50;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 29696 AAAAAAAAAAAAAAAAAAGAAAT 29723
      ||||||||||||||||||||
Db    44 AAAAAAAAAAAAAAAAAAGAAAT 17

```

Search completed: March 30, 2000, 23:51:03
 Job time: 158525 sec



1
2
3

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 30, 2000, 13:20:07 ; Search time 8198.72 Seconds
(without alignments)
13816.022 Million cell updates/sec

Title: US-08-852-495c-1_COPY_115000_145000
Perfect score: 30001
Sequence: 1 AAATCTCCATTTCGTATA.....AGAAATAACCTGATTAAA 30001

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 4538634 seqs, 1887831982 residues

Word size : 0

Total number of hits satisfying chosen parameters: 130662

Minimum DB seq length: 8

Maximum DB seq length: 105

Post-processing: Listing first 45 summaries

Database : EST:*

1: em_est1:*
2: em_est2:*
3: em_est3:*
4: em_est4:*
5: em_est5:*
6: em_est6:*
7: em_est7:*
8: em_est8:*
9: em_est9:*
10: em_est10:*
11: em_est11:*
12: em_est12:*
13: em_est13:*
14: em_est14:*
15: em_est15:*
16: em_est16:*
17: em_est17:*
18: em_est18:*
19: em_est19:*
20: gb_est1:*
21: gb_est2:*
22: gb_est3:*
23: gb_est4:*
24: gb_est5:*
25: gb_est6:*
26: gb_est7:*
27: gb_est8:*
28: gb_est9:*
29: gb_est10:*
30: gb_est11:*
31: gb_est12:*
32: gb_est13:*
33: gb_est14:*
34: gb_est15:*
35: gb_est16:*
36: gb_est17:*
37: gb_est18:*
38: gb_est19:*
39: gb_est20:*
40: gb_est21:*
41: gb_est22:*
42: gb_est23:*
43: gb_est24:*

44: gb_est25:*
45: gb_est26:*
46: gb_est27:*
47: gb_est28:*
48: gb_est29:*
49: gb_est30:*
50: gb_est31:*
51: gb_est32:*
52: em_est20:*
53: em_est21:*
54: em_est22:*
55: em_est23:*
56: em_est24:*
57: em_est25:*
58: em_est26:*
59: gb_est33:*
60: gb_est34:*
61: gb_est35:*
62: gb_est36:*
63: gb_est37:*
64: gb_est38:*
65: em_est27:*
66: em_est28:*
67: em_est29:*
68: em_est30:*
69: gb_est39:*
70: gb_est40:*
71: gb_est41:*
72: gb_est42:*
73: gb_est43:*
74: gb_est44:*
75: em_est31:*
76: em_est32:*
77: em_est33:*
78: em_est34:*
79: gb_est45:*
80: gb_est46:*
81: gb_est47:*
82: gb_est48:*
83: em_est35:*
84: em_est36:*
85: em_est37:*
86: em_est38:*
87: gb_est49:*
88: gb_est50:*
89: gb_est51:*
90: gb_est52:*
91: gb_est53:*
92: em_est39:*
93: em_est40:*
94: em_est41:*
95: em_est42:*
96: em_est43:*
97: em_est44:*
98: em_est45:*
99: gb_est54:*
100: gb_est55:*
101: em_est46:*
102: gb_est56:*
103: gb_est57:*
104: gb_est58:*
105: gb_est59:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
c 1	44	0.1	101	100	A0280224

A0280224 CITBI-EI-

2 42 0.1 93 43 A1168167
 3 41 0.1 95 20 T52775
 4 39 0.1 97 39 AA837701
 5 38 0.1 102 100 A0319270
 6 38 0.1 102 36 AA654562
 7 38 0.1 103 104 A0584425
 8 37 0.1 82 34 AA458985
 9 37 0.1 82 34 AA458985
 10 37 0.1 103 85 A0570476
 11 37 0.1 103 81 B90619
 12 36 0.1 102 104 A0535244
 13 36 0.1 82 104 A0535244
 14 36 0.1 82 104 A0535244
 15 35 0.1 65 81 AA644896
 16 35 0.1 65 81 B36140
 17 35 0.1 101 100 A0278892
 18 35 0.1 105 100 A0278892
 19 34 0.1 50 39 AA665303
 20 34 0.1 65 25 AA665303
 21 34 0.1 101 33 AA381369
 22 33 0.1 104 29 AA128957
 23 33 0.1 91 38 AA670764
 24 33 0.1 95 37 AA670764
 25 33 0.1 101 21 R1703571
 26 33 0.1 101 25 N80946
 27 33 0.1 102 81 B48088
 28 33 0.1 103 104 A0541035
 29 32 0.1 105 28 A078003
 30 32 0.1 72 24 H92864
 31 32 0.1 31 31 AA281256
 32 32 0.1 86 36 AA650598
 33 32 0.1 102 81 B62983
 34 32 0.1 103 38 AA807640
 35 32 0.1 104 50 A1675576
 36 31 0.1 52 20 D25845
 37 31 0.1 37 43 A1206968
 38 31 0.1 84 34 AA501753
 39 31 0.1 95 31 AA291270
 40 31 0.1 100 35 T52775
 41 31 0.1 101 35 AA564832
 42 31 0.1 101 35 AA583697
 43 31 0.1 101 39 AA835205
 44 31 0.1 103 74 N39933
 45 31 0.1 103 74 AA176421
 31 103 81 B48914

ALIGNMENTS

A1168167 o009e10.x
 T52775 ya79d08.r1
 AA837701 o06c02.s
 A0319270 RPCI11-98
 AA654562 nt75f10.s
 A0584425 RPCI11-4
 AA458985 zxc8i08.s
 AA584832 zxc8i08.s
 AA570476 n63h12.s
 B90619 CIT-HSP-216
 A0535244 RPCI11-3
 D05989 HUMG501971
 AA644896 aa93e05.r
 B36140 HS-1038-A1-
 A0278892 CITBI-E1-
 A0278892 RPCI11-94
 AA665303 o988a05.s
 N852303 o988a05.s
 AA128957 ES794442
 AA129957 zn68h04.r
 AA670764 ac68i12.s
 AA669571 ac68i12.s
 R17033 y4c48d10.s
 N80946 ya12d06.r2
 B48088 RPCI11-2.r1
 A0541035 RPCI11-3
 A078003 7h12h08.C
 H92864 y19c11.s
 AA281256 z894e01.s
 AA650598 nu76d17.s
 B62983 CIT-HSP-669
 AA807640 nx08b05.s
 A1675576 wc02c03.x
 D25845 HUMG504217
 A1206968 q130g11.x
 AA501753 ng05e01.s
 AA291270 zs18d04.s
 T52775 ya79d08.r1
 AA564832 n122a06.s
 AA583697 nt85f10.s
 AA835205 ak64h01.s
 N39933 yw68g11.r1
 AA176421 QVO-C7001
 B48914 RPCI11-4A12

Query Match 0.1%; Score 44; DB 100; Length 101;
 Best Local Similarity 100.0%; Pred No. 0.00059;
 Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 DB 70 CAGTGGCTCACCGTATCCGACGCTTGGAGCCAGGT 2273
 |||||
 CAGTGGCTCACCGTATCCGACGCTTGGAGCCAGGT 27

RESULT 2
 LOCUS A1168167/c
 DEFINITION Homo sapiens genomic clone 2522N7,
 genomic survey sequence.
 VERSION A0280224
 KEYWORDS GSS.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (bases 1 to 93)
 AUTHORS NCI-CGAP
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT On Jan 19, 1998 this sequence version replaced gi:2282306.
 Contact: Robert Strausberg, Ph.D.
 Tel: (301) 495-1550
 Email: Robert.Strausberg@nih.gov
 This clone is available
 IMAGE Consortium (info@imgc.nih.gov) for further information.
 Insert Length: 1481 bp
 Seq primer: -40m13 fwd: ET 17m
 High quality sequence stop 79.
 Location/Qualifiers
 1..93
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:1565706"
 /clone_1lb="Soares_NSF_F8_9M_OT_PA_P_S1"
 /lab_host="DH10B"
 /note="Organ: pooled; Vector: pUT3D-Pac (Pharmacia) with
 equal amounts of plasmid DNA from five normalized
 libraries were mixed and ss circles were made in vitro.
 a subcloning HAP purification, this DNA was used as tracer in
 PCR-amplified cDNAs from pools of 5,000 clones made from
 the same 5 libraries. The pools consisted of the following

RESULT 1
 LOCUS A0280224/c
 DEFINITION CITBI-E1-2522N7.TP CITBI-E1 Homo sapiens genomic
 genomic survey sequence.
 VERSION A0280224
 KEYWORDS GSS.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (bases 1 to 101)
 AUTHORS Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K.,
 Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M. and
 Venter,J.C.
 TITLE Use of a random human BAC End Sequence Database for Sequence-Ready
 Map Building (1998)
 JOURNAL Other-GSS: CITBI-E1-2522N7.TP
 COMMENT Contact: Mark Adams
 Department of Eukaryotic Genomes
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA

libraries and clones: Soares NBSF pool 1:
309384-310919, 323208-325895 Soares NBSF pool 1:
145032-147335, 147720-148103, 148872-149255, 15002 -
150407, 151176-152327 Soares NBSF-9M pool 1:
758280-760583, 772104-774407 Soares NBSF pool 1:
304776-306311, 320136-322823, 326280-326653 Soares NBSF
pool 1: 723720-726407, 739080-740999 Subtraction by Bento
Soares and M. Fatima Bonaldo."

BASE COUNT 11 a 28 c 22 g 32 t
Query Match 0.1%; Score 42; DB 43; Length 93;
Best Local Similarity 100.0%; Pred. No. 0.0027;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22440 TTGCACTGAGCCGAATCGCCACTGCTCCAGCTGGCT 22481
Db 81 TTGCACTGAGCCGAATCGCCACTGCTCCAGCTGGCT 40

RESULT 3
LOCUS T52775 95 bp mRNA EST 06-FEB-1995
DEFINITION ya79d08.r1 Stratagene ovary (#937217) Homo sapiens CDNA clone
IMAGE:67887 5' similar to contains Alu repetitive element, mRNA
sequence.
T52775
T52775.1 GI:654635
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 95)
AUTHORS Hallier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chappel, B.,
Chusoe, S., Dietrich, N., Dubeque, T., Favell, A., Gish, W.,
Hawkins, M., Hultman, M., Kucaba, T., Lacy, M., Le, M., Le, N.,
Mardis, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L.,
Rohlfing, T., Scheibenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J.,
Trevisan, E., Underwood, K., Wohlmann, P., Waterston, R., Wilson, R.
and Marra, M.
Generation and analysis of 280,000 human expressed sequence tags
Genome Res. 6 (9), 807-828 (1996)
9704478
Other ESTs: ya79d08.s1
COMMENT Contact: Willson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: estewatson.wustl.edu
Insert Size: 1013
High quality sequence stops: 82 Source: IMAGE Consortium, LNL This
clone is available royalty-free through LNL; contact the IMAGE
Consortium (info@image.llnl.gov) for further information.
Insert Length: 1013 Std Error: 0.00
Seq primer: M13RP1
High quality sequence stop: 82.
Location/Qualifiers
1..95
/organism="Homo sapiens"
/db_xref="GDB:504144"
/db_xref="taxon:9606"
/clone="IMAGE:67887"
/clone_lib="Stratagene ovary (#937217)"
/sex="female"
/dev_stage="49 year old"
/lab_host="SOLR cells (kanamycin resistant)"
/note="Organ: ovary; Vector: Bluescript SK; Site: 1; EcorI;
Site: 2; XhoI; Cloned unidirectionally. Primer: Oligo dt.
Total ovary tissue, normal, caucasian. Average insert
size: 0.8 kb; Uni-ZAP XR Vector; -5' adaptor sequence: 5'

GAATTCGACAG 3' -3' adaptor sequence: 5'
CTGAGTGTGTGTGTGTGTGTGTGT 3'

BASE COUNT 19 a 32 c 20 g 24 t
Query Match 0.1%; Score 41; DB 20; Length 95;
Best Local Similarity 100.0%; Pred. No. 0.0056;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25732 AGAGATCACTGACCTGGAGGAGAGGTTGACGTAG 25772
Db 84 AGAGATCACTGACCTGGAGGAGAGGTTGACGTAG 44

RESULT 4
LOCUS AA837701 97 bp mRNA EST 07-APR-1998
DEFINITION oe06c02.s1 NCI CGAP_Ov2 Homo sapiens CDNA clone IMAGE:1385090
similar to contains element PIR5 repetitive element; mRNA
sequence.
AA837701
AA837701.1 GI:2912900
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 97)
AUTHORS NCI CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
COMMENT On Jan 17, 1998 this sequence version replaced gi:1900439.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Christopher A. Moskaluk, M.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: David B. Kitzman, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone Distribution: NCI CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
www-bio.llnl.gov/bdrrp/image/image.html

Insert Length: 451 Std Error: 0.00
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 50.
Location/Qualifiers
1..97
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1385090"
/clone_lib="NCI CGAP_Ov2"
/sex="female"
/tissue_type="ovary"
/lab_host="DH10B"
/note="Vector: pAMP10; mRNA made from invasive ovarian
tumor, CDNA made by oligo-dT priming. Non-directionally
cloned. Size-selected on agarose gel, average insert size
600 bp. Reference: Kitzman et al. (1996) Cancer Research
56:5380-5383."

BASE COUNT 13 a 24 c 29 g 31 t
ORIGIN

Query Match 0.1%; Score 39; DB 39; Length 97;
Best Local Similarity 100.0%; Pred. No. 0.024;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 25734 GAGATCACTGACCTGGAGGAGAGGTTGACGTAG 25772
|||||

```
DB 97 GAGATACCTTGACCTGGAGGAGGAGGTTCATGAG 59

RESULT 5
LOCUS AO319270 102 bp DNA GSS 06-MAY-1999
DEFINITION RPII1-98B22.TJ RPII-11 Homo sapiens genomic clone RPII-11-98B22,
ACCESSION AO319270
VERSION AO319270.1 GI:4052235
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 102)
Adams M.D., Rounsley, S.D., Zhao S., Bass S., Linher, K., Golden, K.,
Berry, K., Granger, D., Suh, E., Mible, C., de Jong, P. and Venter, J.C.
Use of human BAC End sequences for sequence-ready map building
Unpublished (1999)
Other_GSSs: RPII1-98B22.TJ
Contact: Shaying Zhao, William Nierman, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208
Email: hbeetlgr.org
Clones are derived from the human BAC library RPII-11. For BAC
library availability, please contact Pieter de Jong
(pieredejong.med.buifalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buifalo.edu/ordering) or from
Research Genetics (info@resgen.com). BAC end search page:
http://www.tigr.org/tldb/humgen/bac_end_search/bec_end_search.html
Seq primer: SP6
Class: BAC ends.
FEATURES
source
Location/Qualifiers
1..102
/organism="Homo sapiens"
/db_xref="GDB:7537293"
/db_xref="taxon:9606"
/clone="RPII-11-98B22"
/clone_11b="RPII-11"
/sex="Male"
/cell_type="Lymphocytes"
/notes="Vector: pBAC3.6; Site_1: EcoRI; Site_2: EcoRI;
RPII1 Human Male BAC Library"
BASE COUNT 26 a 28 c 31 g 17 t
ORIGIN

Query Match 0.1%; Score 39; DB 100; Length 102;
Best Local Similarity 100.0%; Pred. No. 0.024;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22230 CAGTGCTCACCTGTATCCGACCTTGGAGGCC 2228
|||||
DB 42 CAGTGCTCACCTGTATCCGACCTTGGAGGCC 80

RESULT 6
LOCUS AA654562 102 bp mRNA EST 04-NOV-1997
DEFINITION nt75110.s1 NCI_CGAP_P3 Homo sapiens cDNA clone IMAGE:1204363
similar to contains Alu repetitive element; contains element MER22
repetitive element; mRNA sequence.
ACCESSION AA654562
VERSION AA654562.1 GI:2590716
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 102)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer genome Anatomy Project (CGAP).
JOURNAL Unpublished (1997)
COMMENT tumor gene index
On Sep 12, 1996 this sequence version replaced gi:1393453.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: W. Marston Linehan, M.D., Rodrigo Chuquai,
M.D., Michael Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: David B. Kitzman, Ph.D.
CDNA Library Arrayed by: Genome Systems Inc., Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNT at:
www.bio.llnl.gov/bbrp/image/image.html

FEATURES
source
Location/Qualifiers
1..102
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="1204363"
/clone_11b="NCI_CGAP_P3"
/sex="Male"
/dev_stage="45 years old"
/lab_host="DH10B"
/notes="Vector: pAMP10; Site_1: NotI; Site_2: EcoRI; 1st
strand cDNA was primed with oligo(dT)17 on 50 ng of
DNase-treated, total cellular RNA obtained from
5,000-10,000 microdissected cells
histologically determined to be fully malignant prostate
cancer cells. Double-stranded cDNA was ligated to EcoRI
adaptors, 5 cycles of PCR applied to the cDNA with an
adaptor-specific primer, and the resulting PCR product
subcloned into pAMP10 by the UDG-cloning method (Life
Technologies). Average insert size is 600 bp. NOTE: Not
directionally cloned. This library was constructed by
David Kitzman."
BASE COUNT 22 a 32 c 27 g 21 t
ORIGIN

Query Match 0.1%; Score 38; DB 36; Length 102;
Best Local Similarity 100.0%; Pred. No. 0.05;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10130 GGAAGTTGAGACCCAGCTGGCCACATGTGTAACCTC 10167
|||||
DB 42 GGAAGTTGAGACCCAGCTGGCCACATGTGTAACCTC 5

RESULT 7
LOCUS AO584425 103 bp DNA GSS 07-JUN-1999
DEFINITION RPII-11-45B12.TJ RPII-11 Homo sapiens genomic clone RPII-11-45B12,
ACCESSION AO584425
VERSION AO584425.1 GI:5011535
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 103)
Zhao, S., Adams M.D., Nierman, W., Malek, J., de Jong, P. and
Venter, J.C.
Use of BAC End sequences from library RPII-11 for Sequence-Ready
Map Building
Unpublished (1997)
Other_GSSs: RPII-11-45B12.TJ
Contact: Shaying Zhao, William Nierman, Mark Adams
```

Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208

Clones are delivered from the human BAC library RCGI-11. For BAC library availability, please contact pierer de jong (p.ieterdejong.med.buffalo.edu). Clones may be purchased from BACpac Resources (<http://Bacpac.med.buffalo.edu/ordering>) or from Research Genet cs (<http://www.researchgenet.com>). BAC end search page: http://www.tigr.org/tdb/humgen/bac_and_search/bac_end_search.html

Seq primer: Sp6

Class: BAC ends.

FEATURES

```

source      1. .103
            /organism="Homo sapiens"
            /db_xref="GDB:767573"
            /db_xref="taxon:9606"
            /clone="RPC1-11-458L2"
            /clone_1lb="RPC1-11"
            /sex="Male"
            /cell_type="Lymphocytes"
            /note="Vector: pBAC3.6; Site.1: EcoRI; Site.2: EcoRI;
            RPc11 Human Male BAC Library"
BASE COUNT 23 a      27 c      31 g      22 t
ORIGIN

```

Query Match	0.18;	Score 38;	DB 104;	Length 103;
Best Local Similarity	100.0%;	Pred. NO. 0.05;		
Matches 38;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

Qy 16505 TGATCTGCCCACTTGGCCTCCCAAGTGTGGATAA 16542

Db 56 TGATCTGCCACCTTGGCTCCCAAGTGTGGATA 19

RESULT 8

LOCUS	AA458985	82 bp	mRNA	EST	09-JUN-1997
DEFINITION	X288h08.s1 Soares Ovary tumor NbHOT Homo sapiens cDNA clone				
IMAGE	810879 3' similar to contains LTRs.t2 LTR5 repetitive element				
;;	mRNA sequence.				

ACCESSION	AA458985
VERSION	AA458985.1
FILE	GI:2183892

SOURCE

ORGANISM

REFERENCE
AUTHORS

Eumariota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Dutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 82)
Hillier, L., Allen, M., Bowles, L., Duñague, T., Geisel, G., Jost

Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F., Theising, B., White, Y., Wylie, T., Waterston, R. and Wilson, R.

TITLE	WashU-Merck EST Project 1997
JOURNAL	Unpublished (1997)
COMMENT	On Sep 12, 1996 this sequence version replaced gi:1395423

Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810

Email: estewatson.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
Seq primer: -4lnl3 fwd. ET from AmerSham
High quality sequence stop: 70.

FEATURES

```
1. .82
/organism="Homo sapiens"
/db_xref="GDB:6041714"
```

```

/db_xref="taxon:9606"
/clone_image:810879"
/clone_lib="Soares ovary tumor NBHOT"
/sex="Female"
/tissue_type="Ovarian tumor"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: ovary; Vector: pT73D (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCATCTGTAAGTGGAGCGCGCGGGTTTTTTTTTTTTTTT 3'], double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT73 vector (Pharmacia). Library constructed by Bento Soares and M.Fatima Bonaldo."
```

```
OY 19894 CAAGGCAGAGAATTTTCTTAGTCACGAAACAAAATG 19930C
      |||||
Db   9 CAAGGCAGAGAATTTTCTTAGTCACGAAACAAAATG 45
```

RESULT 9

LOCUS	85 bp	MRNA	EST	13-MAY-1999
DEFINITION	F26823	HSPD14446 HM3 Homo sapiens	CDNA clone s4000060510, mRNA sequence	

VERSION F26823.1 GI:4812449

SOURCE human.

ORGANISM

REFERENCE

AUTHORS

TITLE

3'-end-specific cDNA library of human skeletal muscle by DNA sequencing and filter hybridization
Genome Res. 6 (1), 35-42 (1996)

COMMENT

University of Padua
Via Trieste 75, 35121 Padua, Italy
ABI Chromatograms and other information are available on WWW at
<http://grup.bio.unipd.it>.

Source

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="s4000060G10"
/clone_1fb="HM3"
/sex="female"
/tissue-type="pectoral muscle (after mastectomy)"
/mote="Vector: pCDNA1 (Invitrogen); Site_1: BstXI;
Site_2: NotI; The library was constructed by G.
Sanfanchi. This library is not subtracted nor normalized.

```

The first strand cDNA was primed with a biotinylated 0.190- μ g NotI primer (5'-biotin-AACCGGCGTCCAGCGCCGCTTTTTTTTTTTTTTTT-3'). The cDNA was sonicated and size-selected in the range 350-550 bp. The 3' specific fragments were selected by streptavidin coated magnetic beads, ligated to non-palindromic BstXI adapters, NotI digested and directionally cloned into BstXI-NotI cut pCDM11 vector.*

JOURNAL COMMENT
Unpublished (1997)
Contact: Shaying Zhao, William Niernman, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208
Email: hbeef@igf.org

Clones are derived from the human BAC library RPC1-11. For BAC library availability, please contact Pieter de Jong (pieter@ejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from Research Genet cs (info@resgen.com). BAC end search page: http://www.tigr.org/tdb/hungen/bac_end_search/bac_end_search.html.
Seq primer: T7
Class: BAC ends

FEATURES

source

1. 103
/organism="Homo sapiens"
/db_xref="GDB:7621533"
/db_xref="taxon:9606"
/clone="RPC1-11-31H22"
/clone_lib="RPC1-11"
/sex="Male"
/cell_type="Lymphocytes"
/note="Vector: pBAC3.6, site_1: EcoRI, site_2: EcoRI;
RPC111 Human Male BAC library"
BASE COUNT 31 a 27 c 27 g 18 t
ORIGIN

Query Match 0.1%; Score 37; DB 104; Length 103;
Best Local Similarity 100.0%; Pred. No. 0.1;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7322 AGAGACGGGTTTCACCATGTTGGCCAGCGTGCTC 7358
|||||
Db 87 AGAGACGGGTTTCACCATGTTGGCCAGCGTGCTC 51
|||||

RESULT 13 87 bp mRNA EST 30-JUL-1996
D20989/c LOCUS HUMG501971 Human promyelocyte Homo sapiens cDNA clone mp0383 3',
DEFINITION mRNA sequence.
ACCESSION D20989
VERSION D20989.1 GI:504809
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 87)
Okubo,K., Fukushima,A., Yoshii,J., Miyama,T., Kojima,Y.,
Yoshinari,H., Arimoto,J. and Matsubara,K.
Gene expression of human promyelocytic cell line HL60 before and
after induction of differentiation. A new application of 3'directed
cDNA sequencing
Unpublished (1993)
CONTACT: Okubo,K., Fukushima,A., Yoshii,J., Miyama,T., Kojima,Y.,
Yoshinari,H., Arimoto,J. and Matsubara,K.
Institute for Molecular and Cellular Biology
Osaka University
3-1 Yamada-oka,Suita,Osaka 565,Japan.

JOURNAL COMMENT

FEATURES

source

1. 87
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="mp0383"
/clone_lib="Human promyelocyte"
/note="Female, adult, cell_line = HL60, cell_type =
promyelocyte."
BASE COUNT 13 a 23 c 17 g 34 t

ORIGIN

Query Match 0.1%; Score 36; DB 20; Length 87;
Best Local Similarity 100.0%; Pred. No. 0.24;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22232 GTGGCTCACACCTGTATCCACACTTTGGGAGGC 22267
|||||
Db 51 GTGGCTCACACCTGTATCCACACTTTGGGAGGC 16
|||||

RESULT 14 97 bp mRNA EST 05-MAR-1998
AA464896/c LOCUS aas9e05.r1 Striatogene fetal retina 937202 Homo sapiens cDNA clone
DEFINITION IMAGE:838880 5' similar to contains Alu repetitive element; mRNA
sequence.
ACCESSION AA464896
VERSION AA464896
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 97)
Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
Kritman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M.,
Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F.,
Theising,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.
WashU-NCI human EST Project
Unpublished (1997)
On May 5, 1995 this sequence version replaced gi:798258.
CONTACT: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.edu

AUTHORS

TITLE JOURNAL
COMMENT On May 5, 1995 this sequence version replaced gi:798258.
CONTACT: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.edu

This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert length: 758 Std Error: 0.00
Seq primer: -28m3 rev1 ET from Amersham
High quality sequence stop: 97.
Location/Qualifiers
1. 97
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:838880"
/clone_lib="Striatogene fetal retina 937202"
/sex="mixed"
/lab_host="SOLR (kanamycin resistant)"
/note="Vector: pBluescript SK-; Site_1: EcoRI, site_2:
XhoI; Cloned unidirectionally. Primer: Oligo dT. Pooled
retinal tissue. Average insert size: 1.0 kb; Uni-ZAP XR
Vector: -5' adaptor sequence: 5' GATTGCGACGCG 3' -3'
adaptor sequence: 5' CTCGAGTTTCTTTTCTTTT 3'

FEATURES

source

1. 97
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:838880"
/clone_lib="Striatogene fetal retina 937202"
/sex="mixed"
/lab_host="SOLR (kanamycin resistant)"
/note="Vector: pBluescript SK-; Site_1: EcoRI, site_2:
XhoI; Cloned unidirectionally. Primer: Oligo dT. Pooled
retinal tissue. Average insert size: 1.0 kb; Uni-ZAP XR
Vector: -5' adaptor sequence: 5' GATTGCGACGCG 3' -3'
adaptor sequence: 5' CTCGAGTTTCTTTTCTTTT 3'

BASE COUNT

17 a 25 c 26 g 29 t

Query Match 0.1%; Score 36; DB 34; Length 97;
Best Local Similarity 100.0%; Pred. No. 0.23;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22374 GTGCTGTATCCACACTACTCAGAGCGCTGAGCA 22409
|||||
Db 38 GTGCTGTATCCACACTACTCAGAGCGCTGAGCA 3
|||||

RESULT 15 65 bp DNA GSS 17-OCT-1997
B36140/c LOCUS B36140 65 bp DNA GSS 17-OCT-1997

Mon Apr 3 08:24:05 2000

us-08-852-495c-1_copy_115000_145000.rst

Page 8

DEFINITION HS-1038-A1-D06-MR.abi CIT Human Genomic Sperm Library C Homo
sapiens genomic clone Plate-CT 820 Col-11 Row-G, genomic survey
sequence.
ACCESSION B36140
VERSION B36140.1 GI:2535509
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 65)
AUTHORS Mahairas,G.G., Zackrone,K.D., Smith,T., Tipton,S., Schmidt,S.,
Tsalicoff,R., Adajian,C., Blanchard,A., West,A. and Hood,L.E.
TITLE Construction of a Characterized Clone Resource for Genomic
Sequencing: Generation and Preliminary Analysis of 20,000 Sequence
Tagged Connectors
JOURNAL Unpublished (1997)
COMMENT Contact: Mahairas GG, Zackrone KD, Hood L
University of Washington
Seattle, WA 98195, USA
Tel: (206) 616-8744
Fax: (206) 685-7301
Email: kzackron@u.washington.edu
Sequence Tagged Connector
Plate: CT 820 row: G column: 11
Class: BAC ends
High quality sequence stop: 65.
FEATURES
source
1
65
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="Plate-CT 820 Col-11 Row-G"
/clone_lib="CIT Human Genomic Sperm Library C"
/sex="M"
/note="Organ: sperm; Vector: pBelBAC11; BAC Clones in
E-Coli DH10B"
BASE COUNT 25 a 18 c 10 g 12 t
ORIGIN
Query Match 0.1%; Score 35; DB 81; Length 65;
Best local Similarity 100.0%; Pred. No. 0.59;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 16455 AGACGGGGTTTCACCATGTTGTCAGGCTGCTC 16489
|||||
DB 35 AGACGGGGTTTCACCATGTTGTCAGGCTGCTC 1
|||||

Search completed: March 30, 2000, 19:57:22
Job time: 146496 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 30, 2000, 23:41:00 ; Search time 13413 Seconds

(without alignments)
-6791.518 Million cell updates/sec

Title: US-08-852-495C-1_COPY_140000_170000

Perfect score: 30001

Sequence: 1 CCCCTCAATCCCATATGCAC.....TACATATTATATATGTA 30001

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 821193 seqs, -1518192014 residues

Word size : 0

Total number of hits satisfying chosen parameters: 345848

Minimum DB seq length: 8

Maximum DB seq length: 105

Post-processing: Listing first 45 summaries

Database :

GenDb1: *
1: gb_bal: *
2: gb_ba2: *
3: gb_om: *
4: gb_ov: *
5: gb_pat: *
6: gb_ph: *
7: gb_p11: *
8: gb_p12: *
9: gb_p13: *
10: gb_p14: *
11: gb_p15: *
12: gb_p16: *
13: gb_p17: *
14: gb_p18: *
15: gb_p19: *
16: gb_p20: *
17: gb_p21: *
18: gb_p22: *
19: gb_p23: *
20: gb_p24: *
21: gb_p25: *
22: gb_p26: *
23: gb_p27: *
24: gb_p28: *
25: gb_p29: *
26: gb_p30: *
27: gb_p31: *
28: gb_p32: *
29: gb_p33: *
30: gb_p34: *
31: gb_p35: *
32: gb_p36: *
33: gb_p37: *
34: gb_p38: *
35: gb_p39: *
36: gb_p40: *
37: gb_p41: *
38: gb_p42: *
39: gb_p43: *
40: gb_p44: *
41: gb_p45: *
42: gb_p46: *
43: gb_p47: *

44: gb_htg6: *
45: gb_htg7: *
46: em_htg1: *
47: em_htg2: *
48: em_htg3: *
49: em_hum5: *
50: gb_p13: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	44	0.1	80	9 HUMBRKFAA	M36131 Human alpha
C 2	42	0.1	80	9 HUMBRKFAA	M36132 Human alpha
C 3	40	0.1	40	5 A68622	A68622 Sequence 2
C 4	39	0.1	90	5 HUMLDLRF1	K03555 Human low d
C 5	38	0.1	60	5 AR051487	AR051487 Sequence
C 6	38	0.1	76	5 AR051499	AR051499 Sequence
C 7	38	0.1	80	9 HUMBRKFAA	M36133 Human alpha
C 8	38	0.1	90	9 HUMLDLRF1	M15365 Human low d
C 9	38	0.1	97	9 HUMLDLRF2	M14180 Human low d
C 10	37	0.1	97	9 HUMLDLRF3	M14179 Human fam1
C 11	37	0.1	101	10 S79560	S79560 HX (Inton
C 12	37	0.1	101	10 S79561	S79561 dHX (partl
C 13	36	0.1	40	5 A68621	A68621 Sequence 1
C 14	36	0.1	51	9 HUMDB03M3	D17279 Human HepG2
C 15	36	0.1	103	9 HUMALCE221	M87896 Human carcl
C 16	36	0.1	105	13 G32655	G32655 A009130 Hum
C 17	35	0.1	35	5 A25212	A25212 Inter-Alu S
C 18	35	0.1	35	5 E09140	E09140 Synthetic D
C 19	35	0.1	70	9 HSLAS32	X91547 H. sapiens D
C 20	35	0.1	91	13 HUMUT8164A	L30244 Human STS U
C 21	34	0.1	63	13 HUMUT5302A	L30829 Human STS U
C 22	34	0.1	80	9 HUMBRKFAA	M36132 Human alpha
C 23	34	0.1	80	9 HUMBRKFAA	M36133 Human alpha
C 24	33	0.1	80	9 HUMBRKFAA	M36135 Human alpha
C 25	33	0.1	100	9 HUMGALNSA	D45223 Human GALNS
C 26	32	0.1	51	10 S62605	S62605 C1-Inhibito
C 27	31	0.1	60	13 HUMUT243A	L29936 Human STS U
C 28	31	0.1	67	9 HUMALFAND	L36843 Homo sapien
C 29	31	0.1	77	13 HUMUT6154A	L30742 Human STS U
C 30	31	0.1	77	13 HUMUT6154A	L30742 Human STS U
C 31	31	0.1	90	9 HUMLDLRF1	K03556 Human low d
C 32	31	0.1	97	9 HUMLDLRF1	M14178 Human low d
C 33	31	0.1	104	9 HUMALCE272	M87899 Human carcl
C 34	30	0.1	50	10 S62604	S62604 C1-Inhibito
C 35	30	0.1	63	13 HUMUT5302A	L30829 Human STS U
C 36	30	0.1	91	13 HUMUT8164A	L30244 Human STS U
C 37	30	0.1	95	4 GGER10H7	X78616 G. gallus ge
C 38	29	0.1	99	13 HUMUT7692A	L30306 Human STS U
C 39	29	0.1	30	5 AR051440	AR051440 Sequence
C 40	29	0.1	84	5 AR051521	AR051521 Sequence
C 41	28	0.1	28	5 A49272	A49272 Sequence 2
C 42	28	0.1	50	5 I23510	I23510 Sequence 5
C 43	28	0.1	50	5 I28359	I28359 Sequence 5
C 44	28	0.1	50	5 I28514	I28514 Sequence 5
C 45	28	0.1	50	5 I41125	I41125 Sequence 5

ALIGNMENTS

RESULT 1
HUMBRKFAA/c
LOCUS HUMBRKFAA 80 bp DNA
DEFINITION Human alpha-galactosidase breakpoint region.
ACCESSION M36131
VERSION M36131.1 GI:179541 27-APR-1993

```
KEYWORDS      Fabry disease; breakpoint junction; glycosphingolipid catabolism.
SOURCE        Human DNA.
ORGANISM      Homo sapiens
REFERENCE     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
AUTHORS      Eutheria; Primates; Catarrhini; Homiidae; Homo.
TITLE        1 (bases 1 to 80)
JOURNAL      Kornreich, R., Bishop, D.F. and Desnick, R.J.
MEDLINE      Alpha-galactosidase A gene rearrangements causing Fabry disease:
FEATURES      Identification of short direct repeats at breakpoints in an
              Alu-rich gene
              J. Biol. Chem. 265, 9319-9326 (1990)
              90264427
              Location/Qualifiers
              source
                1..80
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
              BASE COUNT      17 a      29 c      16 g      18 t
              ORIGIN
Query Match      0.1%; Score 44; DB 9; Length 80;
Best Local Similarity 100.0%; Pred. No. 2.7e-10;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 13518 GTGGTGCATGCTGTATCCAGCAGCTTTGGAGGCCAAG 13561
|||||
Db 60 GTGGTGCATGCTGTATCCAGCAGCTTTGGAGGCCAAG 17

RESULT 2
LOCUS      HMBKRFAB 80 bp DNA PRI 27-APR-1993
DEFINITION Human alpha-galactosidase breakpoint family F.
ACCESSION  M36132
VERSION     M36132.1 GI:179542
KEYWORDS    Fabry disease; breakpoint junction; glycosphingolipid catabolism.
SOURCE      Human DNA.
ORGANISM    Homo sapiens
REFERENCE   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
AUTHORS     Eutheria; Primates; Catarrhini; Homiidae; Homo.
TITLE       1 (bases 1 to 80)
JOURNAL     Kornreich, R., Bishop, D.F. and Desnick, R.J.
MEDLINE     Alpha-galactosidase A gene rearrangements causing Fabry disease:
FEATURES     Identification of short direct repeats at breakpoints in an
              Alu-rich gene
              J. Biol. Chem. 265, 9319-9326 (1990)
              90264427
              Location/Qualifiers
              source
                1..80
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
              BASE COUNT      17 a      29 c      16 g      18 t
              ORIGIN
Query Match      0.1%; Score 42; DB 9; Length 80;
Best Local Similarity 100.0%; Pred. No. 2.7e-09;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 13520 GTGGTGCATGCTGTATCCAGCAGCTTTGGAGGCCAAG 13561
|||||
Db 58 GTGGTGCATGCTGTATCCAGCAGCTTTGGAGGCCAAG 17

RESULT 3
LOCUS      A68622/c 40 bp DNA PAT 06-MAY-1999
DEFINITION Sequence 2 from Patent WO9801573.
ACCESSION  A68622
VERSION     A68622.1 GI:4759649
KEYWORDS    unidentified.
SOURCE      unidentified.
ORGANISM    unidentified.

KEYWORDS      unclassified.
SOURCE        1 (bases 1 to 40)
REFERENCE     Resnick, M.A., Laktionov, V.L., Kornprina, N.Y. and Perkins, E.L.
AUTHORS      TRANSFORMATION-ASSOCIATED RECOMBINATION CLONING
TITLE        Patent: WO 9801573-A 15-JAN-1998;
JOURNAL      US HEALTH (US)
FEATURES      Location/Qualifiers
              source
                1..40
                /organism="unclassified"
                /db_xref="taxon:32644"
              BASE COUNT      9 a      8 c      19 g      4 t
              ORIGIN
Query Match      0.1%; Score 40; DB 5; Length 40;
Best Local Similarity 100.0%; Pred. No. 2.7e-08;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 28376 CCGCTCCCGGTTCAAGCATTCCTGCTCAGCTCC 28415
|||||
Db 40 CCGCTCCCGGTTCAAGCATTCCTGCTCAGCTCC 1

RESULT 4
LOCUS      HMBDLRFL/c 90 bp DNA PRI 11-JAN-1995
DEFINITION Human low density lipoprotein receptor intron A Alu repeat.
ACCESSION  K03555
VERSION     K03555.1 GI:187104
KEYWORDS    Alu repeat; low density lipoprotein receptor-1; repeat region.
SOURCE      Human DNA.
ORGANISM    Homo sapiens
REFERENCE   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
AUTHORS     Eutheria; Primates; Catarrhini; Homiidae; Homo.
TITLE       1 (bases 1 to 90)
JOURNAL     Lehman, M.A., Goldstein, J.L., Russell, D.W. and Brown, M.S.
MEDLINE     Duplication of seven exons in LDL receptor gene caused by Alu-Alu
FEATURES     recombination in a subject with familial hypercholesterolemia
              Cell 48 (5), 827-835 (1987)
              87131094
              Location/Qualifiers
              source
                1..90
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /map="19p13.3"
                /gene="LDLR"
                /note="hgml; 240-119-362"
                /db_xref="taxon:9606"
              BASE COUNT      17 a      30 c      24 g      19 t
              ORIGIN      Chromosome 19p13.2-p13.1.
Query Match      0.1%; Score 39; DB 9; Length 90;
Best Local Similarity 100.0%; Pred. No. 8.2e-08;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 13519 TGGTGCATGCTGTATCCAGCAGCTTTGGAGGCC 13557
|||||
Db 80 TGGTGCATGCTGTATCCAGCAGCTTTGGAGGCC 42

RESULT 5
LOCUS      ARO51487 60 bp DNA PAT 29-SEP-1999
DEFINITION Sequence 57 from patent US 5830670.
ACCESSION  ARO51487
VERSION     ARO51487.1 GI:5974851
KEYWORDS    Unknown.
SOURCE      Unknown.
ORGANISM    Unknown.
REFERENCE     unclassified.
              1 (bases 1 to 60)
```



```
AUTHORS de la Monte,S. and Wands,J.R.
TITLE Neural thread protein gene expression and detection of Alzheimer's
JOURNAL Patent: US 5830670-A 57 03-NOV-1998;
FEATURES Location/Qualifiers
SOURCE 1..60
BASE COUNT 12 a 14 c 15 g 19 t
ORIGIN

Query Match 0.1%; Score 38; DB 5; Length 60;
Best Local Similarity 100.0%; Pred. No. 2.6e-07;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9496 GCTAATTTTGTATTTAGTAGAGACAGGGTTTCACC 9533
Db 11 GCTAATTTTGTATTTAGTAGAGACAGGGTTTCACC 48

RESULT 6
AR051499 76 bp DNA PAT 29-SEP-1999
LOCUS AR051499 Sequence 69 from patent US 5830670.
ACCESSION AR051499
VERSION AR051499.1 GI:5974863
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 76)
AUTHORS de la Monte,S. and Wands,J.R.
TITLE Neural thread protein gene expression and detection of Alzheimer's
JOURNAL Patent: US 5830670-A 69 03-NOV-1998;
FEATURES Location/Qualifiers
SOURCE 1..76
BASE COUNT 15 a 19 c 18 g 24 t
ORIGIN

Query Match 0.1%; Score 38; DB 5; Length 76;
Best Local Similarity 100.0%; Pred. No. 2.6e-07;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9496 GCTAATTTTGTATTTAGTAGAGACAGGGTTTCACC 9533
Db 11 GCTAATTTTGTATTTAGTAGAGACAGGGTTTCACC 48

RESULT 7
HUMBRKAC/c 80 bp DNA PRI 27-APR-1993
LOCUS HUMBRKAC Human alpha-galactosidase breakpoint region.
DEFINITION M36133
ACCESSION M36133.1 GI:179543
VERSION M36133.1
KEYWORDS Fabry disease; breakpoint junction; glycosphingolipid catabolism.
SOURCE Human DNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 80)
AUTHORS Kornreich,R., Bishop,D.F. and Desnick,R.J.
TITLE Alpha-galactosidase A gene rearrangements causing Fabry disease:
identification of short direct repeats at breakpoints in an
Alu-rich gene
JOURNAL J. Biol. Chem. 265, 9319-9326 (1990)
MEDLINE 90264427
FEATURES Location/Qualifiers
SOURCE 1..80
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 18 a 28 c 17 g 17 t
ORIGIN

Query Match 0.1%; Score 38; DB 9; Length 80;
Best Local Similarity 100.0%; Pred. No. 2.6e-07;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13520 GGTGGCTCATGCTGTAAATCCACAGCATTGGAGGCC 13557
Db 58 GGTGGCTCATGCTGTAAATCCACAGCATTGGAGGCC 21

RESULT 8
HUMDLRAC/c 90 bp DNA PRI 11-JAN-1995
LOCUS HUMDLRAC Human low density lipoprotein receptor mutant gene recombination
DEFINITION site.
ACCESSION M15365
VERSION M15365.1 GI:187107
KEYWORDS Alu repeat; LDL receptor; cell surface protein; crossover;
recombination.
SOURCE Human (FH 295) fibroblast DNA, clone p295.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 90)
AUTHORS Lehman,M.A., Goldstein,J.L., Russell,D.W. and Brown,M.S.
TITLE Duplication of seven exons in LDL receptor gene caused by Alu-Alu
JOURNAL recombination in a subject with familial hypercholesterolemia
MEDLINE 87131094
COMMENT Clean copy of sequence kindly provided by M. Lehman (27-APR-1987).
Individual FH 295 carries two mutant LDL receptor alleles. The
allele below was inherited from the father and includes a
duplication of exons 2 through 8 that was presumably created by
unequal chromosomal crossover involving Alu repeats in introns 1
and 8. The exact site of recombination cannot be determined
because the sequences of the Alu repeats in introns 1 and 8 of
normal alleles are identical over a span of 26 nucleotides at the
recombination site.
FEATURES Location/Qualifiers
SOURCE 1..90
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="19p13.3"
1..43
/gene="LDLR"
<1..43
/gene="LDLR"
/note="LDLR intron 8; G00-119-362"
43..68
/organism="Homo sapiens"
intron 68..>90
/note="LDLR duplicated intron 1 (no splice consensus at
68); putative; does not fit consensus"
BASE COUNT 18 a 33 c 19 g 20 t
ORIGIN 1 bp upstream of NheI site; chromosome 19p13.2-p13.1.

Query Match 0.1%; Score 38; DB 9; Length 90;
Best Local Similarity 100.0%; Pred. No. 2.6e-07;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13519 TGGTGCTCATGCTGTAAATCCACAGCATTGGAGGCC 13556
Db 80 TGGTGCTCATGCTGTAAATCCACAGCATTGGAGGCC 43

RESULT 9
HUMDLRAC/c 97 bp DNA PRI 07-JAN-1995
LOCUS HUMDLRAC Human low density lipoprotein receptor gene (LDLR), intron 5
DEFINITION
```

```
(partial).
ACCESSION M14180
VERSION M14180.1 GI:187098
KEYWORDS low density lipoprotein receptor-1.
SEGMENT 2 of 2
SOURCE Human white blood cell DNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
AUTHORS Hobbs, H.H., Brown, M.S., Goldstein, J.L. and Russell, D.W.
TITLE Deletion of exon encoding cysteine-rich repeat of low density
lipoprotein receptor alters its binding specificity in a subject
with familial hypercholesterolemia
J. Biol. Chem. 261 (28), 13114-13120 (1986)
JOURNAL MEDLINE 87008518
COMMENT Analysis of the LDL-receptor gene of a patient with familial
hypercholesterolemia (FH) revealed the deletion of exon 5 resulting
from a homologous recombination between repetitive Alu sequences of
intron 4 and intron 5.
FEATURES
source 1..97
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="19p13.3"
join(M14178..11..97,1..97)
/gene="LDLR"
<1..>97
/note="LDLR"
/note="LDLR intron E; G00-119-362"
misc_feature 42..72
/gene="LDLR"
/note="deletion target sequence"
BASE COUNT 19 a 34 c 25 g 19 t
ORIGIN Chromosome 19p13.2-p13.1.

Query Match 0.1%; Score 38; DB 9; Length 97;
Best Local Similarity 100.0%; Pred. No. 2.6e-07;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13529 TGCGTGTATCCAGCACTTGGAGCCAGCAGGC 13566
|||||
Db 79 TGCGTGTATCCAGCACTTGGAGCCAGCAGGC 42

RESULT 10
HMDLDRJ/c 97 bp DNA PRI 27-APR-1993
LOCUS Human familial hypercholesterolemia 626-a gene with a deletion of
DEFINITION exon 5.
ACCESSION M14179
VERSION M14179.1 GI:187101
KEYWORDS low density lipoprotein receptor-1; very low density lipoprotein.
SOURCE Human fibroblast DNA, from patient FH-626.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
AUTHORS Hobbs, H.H., Brown, M.S., Goldstein, J.L. and Russell, D.W.
TITLE Deletion of exon encoding cysteine-rich repeat of low density
lipoprotein receptor alters its binding specificity in a subject
with familial hypercholesterolemia
J. Biol. Chem. 261, 13114-13120 (1986)
JOURNAL MEDLINE 87008518
COMMENT Analysis of the LDL-receptor gene of a patient with familial
hypercholesterolemia (FH) revealed the deletion of exon 5 resulting
from a homologous recombination between repetitive Alu sequences of
intron 4 and intron 5.
FEATURES
source 1..97
/organism="Homo sapiens"
/db_xref="taxon:9606"

misc_feature 42..72
/note="deletion target sequence"
BASE COUNT 22 a 34 c 23 g 18 t
ORIGIN Chromosome 19p13.2-p13.1.

Query Match 0.1%; Score 37; DB 9; Length 97;
Best Local Similarity 100.0%; Pred. No. 8e-07;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13529 TGCGTGTATCCAGCACTTGGAGCCAGCAGGC 13565
|||||
Db 79 TGCGTGTATCCAGCACTTGGAGCCAGCAGGC 43

RESULT 11
S79560/c 101 bp DNA PRI 27-JAN-1996
LOCUS HRX (intron 6) [human, genomic, 101 nt].
DEFINITION HRX (intron 6) [human, genomic, 101 nt].
ACCESSION S79560
VERSION S79560.1 GI:1168041
KEYWORDS human.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
AUTHORS Bernard, O.A., Romana, S.P., Schichman, S.A., Mauchauffe, M.,
Jonveaux, P. and Berger, R.
TITLE Partial duplication of HRX in acute leukemia with trisomy 11
MEDLINE 95387650
REMARK GenBank staff at the National Library of Medicine created this
entry [NCBI 94bbsg 170933] from the original journal article.
Map location: 11q23.
Location/Qualifiers
source 1..101
/organism="Homo sapiens"
/db_xref="taxon:9606"
1..101
/partial
/gene="HRX"
BASE COUNT 27 a 21 c 28 g 25 t
ORIGIN

Query Match 0.1%; Score 37; DB 10; Length 101;
Best Local Similarity 100.0%; Pred. No. 8e-07;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9514 AGTAGACAGCGGTTCCACCATGTTGGCCAGGCTGT 9550
|||||
Db 50 AGTAGACAGCGGTTCCACCATGTTGGCCAGGCTGT 14

RESULT 12
S79561/c 101 bp DNA PRI 27-JAN-1996
LOCUS dHRX (partial genomic duplication startpoint) [human, acute myeloid
DEFINITION leukemia with trisomy 11 syndrome patient J, genomic Mutant, 101
nt].
ACCESSION S79561
VERSION S79561.1 GI:1168042
KEYWORDS human acute myeloid leukemia with trisomy 11 syndrome patient J.
SOURCE human sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
AUTHORS Bernard, O.A., Romana, S.P., Schichman, S.A., Mauchauffe, M.,
Jonveaux, P. and Berger, R.
TITLE Partial duplication of HRX in acute leukemia with trisomy 11
```

JOURNAL
MEDLINE
REMARK
Leukemia 9 (9), 1487-1490 (1995)
95387660
Genbank staff at the National Library of Medicine created this entry [NCBI gidsq 170936] from the original journal article. This sequence comes from Fig. 3.
Map location: 11q23.
Authors indicate breakpoint region at 48-62.
COMMENT
FEATURES
source
1. .101
/organism="Homo sapiens"
/db_xref="taxon:9606"
1. .101
/partial
/gene="dhrx"
BASE COUNT
22 a 22 c 23 g 34 t
ORIGIN

Query Match
Best Local Similarity 100.0%; Score 37; DB 10; Length 101;
Pred. No. 8e-07;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9514 AGTAGAGACGGCTTCCATGCTGGCCAGCGTGT 9550
DB 50 AGTAGAGACAGGGTTCCACATGTTGCCAGCGTGT 14

RESULT 13
LOCUS A68621 40 bp DNA
DEFINITION Sequence 1 from Patent WO9801573.
ACCESSION A68621
VERSION A68621.1 GI:4759648
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 40)
AUTHORS Resnick, M.A., Laktionov, V.L., Kouprina, N.Y. and Perkins, E.L.
TITLE TRANSFORMATION-ASSOCIATED RECOMBINATION CLONING
JOURNAL Patent: WO 9801573-A 15-JAN-1998;
US HEALTH (US)

FEATURES
source
1. .40
Location/Qualifiers
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT
7 a 12 c 13 g 8 t
ORIGIN

Query Match
Best Local Similarity 100.0%; Score 36; DB 5; Length 40;
Pred. No. 2.5e-06;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 24026 GTGGCTACCCCTGATCCACAGACTTTGGAGGC 24061
DB 2 GTGGCTACCGCTGTATCCAGCATTGGAGGC 37

RESULT 14
LOCUS HUMD6B03M3 51 bp mRNA
DEFINITION Human HepG2 3' region MboI cDNA, clone hmd6b03m3.
ACCESSION D17279
VERSION D17279.1 GI:598922
KEYWORDS gene signature.
SOURCE Homo sapiens male cell_line:HepG2 cDNA to mRNA, clone_lib:Kiseru.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 51)
AUTHORS Matoba, R.
TITLE Direct Submission

JOURNAL
MEDLINE
REMARK
Submitted (21-Jul-1993) to the DDBJ/EMBL/Genbank databases. Ryo Matoba, Osaka University, Institute for Molecular and Cellular Bio; 1-3, Yamada-oka, Suita, Osaka 565, Japan
E-mail: matoba@inheri1.lmc.osaka-u.ac.jp;
Tel: 81-6-877-5111 (ex. 3314), Fax: 81-6-877-1922)
2 (bases 1 to 51)
Authors indicate breakpoint region at 48-62.
COMMENT
FEATURES
source
1. .51
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_line="HepG2"
/clone_lib="Kiseru"
/sex="Male"
BASE COUNT
13 a 19 c 12 g 7 t
ORIGIN

Query Match
Best Local Similarity 100.0%; Score 36; DB 9; Length 51;
Pred. No. 2.5e-06;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 19138 CCACGTGCACTCCAGCTGGCGCAGAGGAGACTC 19173
DB 8 CCACGTGCACTCCAGCTGGCGCAGAGGAGACTC 43

RESULT 15
LOCUS HUMALCE221/C 103 bp ss-RNA
DEFINITION Human carcinoma cell-derived Alu RNA transcript, clone CE221.
ACCESSION M87896
VERSION M87896.1 GI:174874
KEYWORDS Alu repeat.
SOURCE Homo sapiens male embryo carcinoma cDNA to other RNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 103)
AUTHORS Simmet, D., Richer, C., Deragon, J.-M. and Labuda, D.
TITLE Alu RNA transcripts in human embryonal carcinoma cells. Model of post-transcriptional selection of master sequences
JOURNAL J. Mol. Biol. (1992) in press
FEATURES
source
1. .103
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_line="HepG2"
/dev_stage="embryo"
/sex="male"
/tissue_type="carcinoma"
BASE COUNT
25 a 27 c 33 g 18 t
ORIGIN

Query Match
Best Local Similarity 100.0%; Score 36; DB 9; Length 103;
Pred. No. 2.5e-06;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 28372 ACCTCGGCTCCCGGGTTCAAGCATTCCTGCT 28407

Mon Apr 3 08:24:06 2000

us-08-852-495c-1_copy_140000_170000.rge

Page 6

Db 69 ACCCTCGGCTCCCGGGTTCAGGAGATCTCTCGCT 34

Search completed: March 31, 2000, 06:36:34
Job time: 184343 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 31, 2000, 00:01:30 ; Search time 600.33 Seconds

(without alignments)
12503.144 Million cell updates/sec

Title: US-08-852-495c-1_COPY_140000_170000
Perfect score: 30001
Sequence: 1 CCGTCCCAATCCATATGCAC.....TACATATTTATATTTATGTA 30001

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 311585 seqs, 125096042 residues

Word size : 0

Total number of hits satisfying chosen parameters: 433264

Minimum DB seq length: 8

Maximum DB seq length: 105

Post-processing: Listing first 45 summaries

Database : N_Geneseq_36.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	40	0.1	40	1	V19045
C 2	36	0.1	40	1	V19044
C 3	35	0.1	35	1	Q27391
C 4	29	0.1	30	1	Q77890
C 5	29	0.1	30	1	T27744
C 6	29	0.1	55	1	T98499
C 7	29	0.1	58	1	T98500
C 8	28	0.1	28	1	T12509
C 9	28	0.1	32	1	Q73570
C 10	28	0.1	32	1	Q73570
C 11	28	0.1	40	1	O55168
C 12	28	0.1	48	1	T04081
C 13	28	0.1	49	1	T02146
C 14	28	0.1	49	1	T88080
C 15	28	0.1	53	1	V54636
C 16	28	0.1	60	1	T69401
C 17	28	0.1	86	1	Q11760
C 18	28	0.1	98	1	X00174
C 19	28	0.1	99	1	X23764
C 20	28	0.1	101	1	T07155
C 21	27	0.1	29	1	V15487
C 22	27	0.1	36	1	V15458
C 23	27	0.1	41	1	V03013
C 24	27	0.1	43	1	V48090
C 25	27	0.1	43	1	V48091
C 26	27	0.1	47	1	T98523
C 27	27	0.1	48	1	Q93860
C 28	27	0.1	50	1	T39286
C 29	27	0.1	50	1	T92389
C 30	27	0.1	50	1	V26819
C 31	27	0.1	50	1	V37933
C 32	27	0.1	50	1	X02643
C 33	27	0.1	50	1	X24942

C 34	27	0.1	51	1	T98497
C 35	27	0.1	52	1	T33764
C 36	27	0.1	52	1	T98502
C 37	27	0.1	52	1	T88005
C 38	27	0.1	52	1	V24462
C 39	27	0.1	52	1	V24019
C 40	27	0.1	52	1	V35496
C 41	27	0.1	52	1	X19061
C 42	27	0.1	53	1	V83890
C 43	27	0.1	54	1	T17031
C 44	27	0.1	54	1	T72959
C 45	27	0.1	55	1	Q36440

ALIGNMENTS

RESULT 1
V19045/c
ID V19045 standard; DNA: 40 BP.
AC V19045;
DT 28-JUL-1998 (first entry)
DE Alu PCR primer 2; amplification: Alu repeat sequence; vector;
KW PCR, primer; amplification: Alu repeat sequence; vector;
OS Synthetic.
OS Saccharomyces sp.
PN MO9801573-Al.
PD 15-JAN-1998.
PF 09-JUL-1996; U11478.
PR (USSH) US DEPT HEALTH & HUMAN SERVICES.
PI Kourina NY, Larionov VL, Perkins EL, Resnick MA;
DR WPI: 98-110234/10.
PT Preparation of yeast artificial chromosomes - by in vivo
recombination using vector comprising yeast centromere, marker,
PT yeast telomere and nucleic acid for recombination
PS Example 1; Page 45; 117pp; English.
CC This is the nucleotide sequence for the PCR primer used in the
CC amplification of the Alu repeat sequence, which is used to
CC demonstrate the processes described in the invention. It involves
CC the creation and use of circular yeast artificial chromosome (YAC)
CC to selectively clone specific nucleic acids from a background of
CC mixed nucleic acids by introducing the vector(s) into E. coli cells.
CC They can be used to rapidly isolate human DNA where only a part of the
CC sequence of DNA is known. Using the methods large fragments of DNA can
CC be easily cloned and analysed.
SQ Sequence 40 BP; 9 A; 8 C; 19 G; 4 T;
Query Match 0.1%; Score 40; DB 1; Length 40;
Best Local Similarity 100.0%; Pred. No. 0.032; 0; Indels 0; Gaps 0;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 28376 CCGCTCCCGGCTCAGCGATTCCTCCGCTCAGCCTCC 28415
DB 40 CCGCTCCCGGCTCAGCGATTCCTCCGCTCAGCCTCC 1
RESULT 2
V19044
ID V19044 standard; DNA: 40 BP.
AC V19044;
DT 28-JUL-1998 (first entry)
DE Alu PCR primer 1.
KW PCR, primer; amplification: Alu repeat sequence; vector;
OS Synthetic.
OS Saccharomyces sp.
PN MO9801573-Al.
PD 15-JAN-1998.
PF 09-JUL-1996; U11478.
PR 09-JUL-1996; MO-011478.

PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
PI KOPRISHA NY LARIONOV VL, FERRINS EL, KESNICK MA;
DR MPI: 36-110234/10.
PT Preparation of yeast artificial chromosomes - by in vivo
PT recombination using vector comprising yeast centromere, marker,
PS yeast telomere and nucleic acid for recombination
PI Example 1; Page 45; 11/99; English.
CC This is the nucleotide sequence for the PCR primer used in the
CC amplification of the Alu repeat sequence, which is used to
CC demonstrate the processes described in the invention. It involves
CC the creation and use of circular yeast artificial chromosome (YAC)
CC to selectively clone specific nucleic acids from a background of
CC mixed nucleic acids by introducing the vector(s) into E. coli cells.
CC They can be used to rapidly isolate human DNA where only a part of the
CC sequence of DNA is known. Using the methods large fragments of DNA can
CC be easily cloned and analysed.
SQ Sequence 40 BP; 7 A; 12 C; 13 G; 8 T;

Query Match 0.1%; Score 36; DB 1; Length 40;
Best Local Similarity 100.0%; Pred. No. 0.36; Mismatches 0; Gaps 0;
Matches 36; Conservative 0; Indels 0;

OY 24026 GTGGCTACGCGCTGTATCCAGCACTTGGAGGC 24061

DB 2 GTGGCTACGCGCTGTATCCAGCACTTGGAGGC 37

RESULT 3
OY 027391/c
ID 027391 standard; DNA; 35 BP.
AC 027391; 27-JAN-1993 (first entry)
DE Inter-Alu specific primer PDJ33.
KW Polymerase chain reaction; PCR; repetitive element; ss.
OS Synthetic.
PN WO9213101-A.
PD 06-AUG-1992.
PF 24-JAN-1992; NL0018.
PR 25-JAN-1991; NL-000132.
PA (INGE-) INGENY BV.
PI Uiterlinden AG, Vijg J;
DR MPI: 92-284683/34.
PT Detection of genetic variation by 2-D electrophoresis of
PT fragments - and hybridisation with labelled probes, carried out
PT on fragments consisting of inter-repeat sequences generated by
PT PCR.
PS Claim 6; Page 6; 31pp; English.
CC Primer PDJ33 is one of several primers which are preferred for use
CC in amplifying inter-Alu regions of DNA. The amplified fragments are
CC then subjected to 2-D electrophoresis on the basis of length and
CC differences in base sequence. The resulting separation pattern is
CC transferred to a filter for screening with a probe. The method can
CC be used to detect genetic variation.
CC See 027389-027404 and 033141-033144.
SQ Sequence 35 BP; 8 A; 10 C; 11 G; 6 T;

Query Match 0.1%; Score 35; DB 1; Length 35;
Best Local Similarity 100.0%; Pred. No. 0.69; Mismatches 0; Gaps 0;
Matches 35; Conservative 0; Indels 0;

OY 24027 TGGCTACGCGCTGTATCCAGCACTTGGAGGC 24061

DB 35 TGGCTACGCGCTGTATCCAGCACTTGGAGGC 1

RESULT 4
ID 077890/c
AC 077890 standard; CDNA; 30 BP.
DE 06-JUL-1995 (first entry)
PI Neutral thread protein AD10-7 CDNA 5' antisense oligonucleotide.

KW Neutral thread protein AD10-7; Alzheimer's; neuroectodermal tumours;
KW malignant astrocytomas; glioblastomas; 5' antisense therapy; ss.
OS Synthetic.
PN WO9423756-A.
PD 27-OCT-1994.
PF 20-APR-1994; US-050359.
PR (GHEO) GEN HOSPITAL CORP.
PI De LA MONTE SM, Wands JR;
DR MPI: 94-341497/42.
PT Detection of neural thread proteins - to detect sporadic and
PT familial Alzheimer's disease, neuroectodermal tumours, malignant
PT astrocytomas and glioblastomas (Eng).
PS Disclosure; Page 48; 18pp; English.
CC 077888-077890 are AD10-7 neutral thread protein (NTP) antisense
CC oligonucleotides, that can be used to down regulate or inhibit
CC the expression of the NTP gene. These oligonucleotides could be
CC used in the treatment of the following conditions Alzheimer's
CC disease, neuroectodermal tumours, malignant astrocytomas and
CC glioblastomas.
SQ Sequence 30 BP; 8 A; 4 C; 14 G; 4 T;

Query Match 0.1%; Score 29; DB 1; Length 30;
Best Local Similarity 100.0%; Pred. No. 27; Mismatches 0; Gaps 0;
Matches 29; Conservative 0; Indels 0;

OY 28388 TTCAAGGATTCCTCGCTGACGCC 28416

DB 30 TTCAAGGATTCCTCGCTGACGCC 2

RESULT 5
ID T27744/c
ID T27744 standard; DNA; 30 BP.
AC T27744; 14-NOV-1996 (first entry)
DE Neutral thread protein antisense sequence.
KW Neutral thread protein; NTP; diagnosis; detection;
KW Alzheimer's disease; neuroectodermal tumour; malignant astrocytoma;
KW monoclonal antibody; binding fragment; ds.
OS Synthetic.
PN WO9615272-A1.
PD 23-MAY-1996.
PF 14-NOV-1995; U17111.
PR 14-NOV-1994; US-340426.
PA (GHEO) GEN HOSPITAL CORP.
PI De LA MONTE S, Wands JR;
DR MPI: 96-259865/26.
PT Detection of neural thread protein in diagnosis of Alzheimer's
PT disease - also NTP DNA and protein sequences used in gene and
PT anti-sense therapy.
PS Disclosure; Page 48; 23pp; English.
CC A method for detecting the presence of neural thread protein (NTP)
CC having a molecular weight of 8, 14, 17, 21, 26 or 42 kD in a human
CC subject comprises (a) contacting a sample from a human subject that
CC is suspected of containing the NTP with at least one molecule
CC capable of binding to the protein, and (b) detecting any of the
CC from an antibody free of natural impurities, a monoclonal antibody
CC or a binding fragment of either of these. The method may be used for
CC diagnosing the presence of Alzheimer's disease, neuroectodermal
CC tumours and a malignant astrocytoma in a human. Expression of NTP
CC nucleic acid may be inhibited using antisense oligonucleotides
CC (See T27739-44).
SQ Sequence 30 BP; 8 A; 4 C; 14 G; 4 T;

Query Match 0.1%; Score 29; DB 1; Length 30;
Best Local Similarity 100.0%; Pred. No. 27; Mismatches 0; Gaps 0;
Matches 29; Conservative 0; Indels 0;

OY 28388 TTCAAGGATTCCTCGCTGACGCC 28416

DB 30 TTCACGAGTCTCTGCGCCACGCTCC 2
|||||
RESULT 6
ID T98499 standard; DNA; 55 BP.
AC T98499;
DT 11-MAR-1998 (first entry)
DE CDNA synthesis primer CDS3.
KW Template switching oligonucleotide; RNA/DNA hybrid; DNA preparation;
KW CDNA synthesis primer; CDNA cloning; CDNA library construction;
KW PCR primer; amplify; ss.
OS Synthetic.
PN MO9724455-A2.
PD 10-JUL-1997.
PF 03-JAN-1997; U00368.
PR 03-JAN-1996; US-582562.
PA (CLON-) CLONTECH LAB INC.
PI Chenchik A, Diatchenko L, Siebert P, Zhu Y;
DR WPI: 97-363690/33.
PT Preparation of cDNA from RNA molecules - by annealing CDNA synthesis
PT primer to RNA, synthesizing DNA and contacting with novel template
PT switching oligo:nucleotide
PS Claim 9; Page 29; 39pp; English.
CC T98497-T98520 represent primers used in the method of the invention. The
CC method of the invention is for preparing DNA complementary to the 5'-end
CC of an RNA molecule. The method comprises annealing a first DNA strand
CC primer to the RNA molecule and synthesizing a first DNA strand
CC complementary to at least a portion of the RNA molecule, and contacting
CC the RNA molecule with a template switching oligonucleotide (TSO) (such
CC as T98426) having a pre-selected nucleotide sequence at its 5'-end and at
CC least 1 riboguanine residue at its 3'-end, where the TSO binds the 5'-end
CC of the RNA molecule and serves as a template for the extension of the
CC 3'-end of the first DNA strand. The method can be used for the synthesis
CC and cloning of full length cDNA, or fragments, that correspond to the
CC complete sequence of the 5'-end of the mRNA molecule. It can be used to
CC construct cDNA libraries from nanogram quantities of total or poly A+
CC RNA. The TSO allows for negative selection against cDNA that are not
CC complementary to the 5'-end of the template RNA, while allowing full
CC length cDNA to be readily selected.
SO Sequence 55 BP; 10 A; 5 C; 7 G; 33 T;
Query Match 0.1%; Score 29; DB 1; Length 55;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 26680 AAAAAAAAAAAAAAAAAAGTA 26708
|||||
DB 51 AAAAAAAAAAAAAAAAAAGTA 23
RESULT 7
ID T98500 standard; DNA; 58 BP.
AC T98500;
DT 11-MAR-1998 (first entry)
DE CDNA synthesis primer Fr-T10NN.
KW Template switching oligonucleotide; RNA/DNA hybrid; DNA preparation;
KW CDNA synthesis primer; CDNA cloning; CDNA library construction;
KW PCR primer; amplify; ss.
OS Synthetic.
PN MO9724455-A2.
PD 10-JUL-1997.
PF 03-JAN-1997; U00368.
PR 03-JAN-1996; US-582562.
PA (CLON-) CLONTECH LAB INC.
PI Chenchik A, Diatchenko L, Siebert P, Zhu Y;
DR WPI: 97-363690/33.
PT Preparation of cDNA from RNA molecules - by annealing CDNA synthesis
PT primer to RNA, synthesizing DNA and contacting with novel template
PT switching oligo:nucleotide

PS Claim 9; Page 30; 39pp; English.
CC T98497-T98520 represent primers used in the method of the invention. The
CC method of the invention is for preparing DNA complementary to the 5'-end
CC of an RNA molecule. The method comprises annealing a first DNA strand
CC primer to the RNA molecule and synthesizing a first DNA strand
CC complementary to at least a portion of the RNA molecule, and contacting
CC the RNA molecule with a template switching oligonucleotide (TSO) (such
CC as T98426) having a pre-selected nucleotide sequence at its 5'-end and at
CC least 1 riboguanine residue at its 3'-end, where the TSO binds the 5'-end
CC of the RNA molecule and serves as a template for the extension of the
CC 3'-end of the first DNA strand. The method can be used for the synthesis
CC and cloning of full length cDNA, or fragments, that correspond to the
CC complete sequence of the 5'-end of the mRNA molecule. It can be used to
CC construct cDNA libraries from nanogram quantities of total or poly A+
CC RNA. The TSO allows for negative selection against cDNA that are not
CC complementary to the 5'-end of the template RNA, while allowing full
CC length cDNA to be readily selected.
SO Sequence 58 BP; 8 A; 5 C; 9 G; 34 T;
Query Match 0.1%; Score 29; DB 1; Length 58;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 26680 AAAAAAAAAAAAAAAAAAGTA 26708
|||||
DB 52 AAAAAAAAAAAAAAAAAAGTA 24
RESULT 8
ID T12509 standard; DNA; 28 BP.
AC T12509;
DT 09-SEP-1996 (first entry)
DE Primer A1u B corresp. to bases 294-267.
KW Primer: PCR; polymerase chain reaction; amplification; A1u repeat; ss;
KW quantitation; internal standard; plasmid; contamination; therapeutic.
OS Synthetic.
PN AT9401830-A.
PD 15-DEC-1995.
PF 26-SEP-1994; 001830.
PR 26-SEP-1994; AT-001830.
PA (IMMO) IMMUNO AG.
PI Dörner F, Falkner F, Haemmerle T, Himmelspach M;
DR WPI: 96-069063/08.
PT Quantifying genomic DNA by amplification of repetitive sequences -
PT in presence of internal standard, then comparing amts of amplified
PT genomic DNA and standard, partic. for quality control of recombinant
PT viral proteins, vaccines, etc.
PS Claim 13; Page 14; 41pp; German.
CC Primers T12508-9 are used to amplify a 146 bp fragment of an A1u repeat
CC sequence in a novel method of quantitating genomic DNA in a sample. The
CC novel method features the addition of a known amount of an internal
CC standard nucleic acid which is different from the test nucleic acid in
CC at least one detectable feature. The template for these primers is the
CC plasmid pAlu-wt which comprises plasmid pCRII contg. nucleotides 148-294
CC of the A1u repeat sequence given in Nucleic Acid Res., 18 (1990) 6793.
CC The method is esp. useful for the detection of contaminating DNA in the
CC manufacture of therapeutic prods.
SO Sequence 28 BP; 5 A; 9 C; 9 G; 5 T;
Query Match 0.1%; Score 28; DB 1; Length 28;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 19152 CTTGGGCGACAGCGAGACTCTCTC 19179
|||||
DB 28 CTTGGGCGACAGCGAGACTCTCTC 1
RESULT 9

```
073570
ID 073570 standard; DNA: 32 BP.
AC 073570;
DE 25-JUN-1995 (first entry)
KW Enhancer element er-3 conserved basepair sequence.
KW Enhancer element; carcinoma; tumor; cancer; SLP1 gene;
KW secretory leukoprotease-inhibitor gene; cyokeratin gene-8;
OS Homo sapiens.
FT Key Location/Qualifiers
FT misc.difference 29
FT /*tag= a
FT /*label= pyrimidine
PN W09421118-A.
PD 29-SEP-1994.
PR 24-MAR-1994; U03197.
PR 24-MAR-1993; US-035435.
PA (UABR-) UAB RES FOUND.
PI Garver RI, Sorscher EJ.
PI WPI: 94-316537/39.
DR DNA construct for treating human carcinoma - includes a
PT cancer-therapeutic gene under the control of a promoter and a gp.
PT of enhancer sequences
PS Claim 1; Fig. 6; 54pp; English.
CC This enhancer element is part of a DNA construct used for treating
CC human carcinoma which contains a cancer therapeutic protein under
CC the control of a promoter and 3 enhancer sequences in a specific 5'-3'
CC order. This enhancer element is derived from the flanking region of
CC the human epithelial cell cyokeratin-8 gene.
SQ Sequence 32 BP; 7 A; 1 C; 8 G; 15 T;
```

```
Query Match 0.1%; Score 28; DB 1; Length 32;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
073570/c
ID 073570 standard; DNA: 32 BP.
AC 073570;
DE 25-JUN-1995 (first entry)
KW Enhancer element er-3 conserved basepair sequence.
KW Enhancer element; carcinoma; tumor; cancer; SLP1 gene;
KW secretory leukoprotease-inhibitor gene; cyokeratin gene-8;
OS Homo sapiens.
FT Key Location/Qualifiers
FT misc.difference 29
FT /*tag= a
FT /*label= pyrimidine
PN W09421118-A.
PD 29-SEP-1994.
PR 24-MAR-1994; U03197.
PR 24-MAR-1993; US-035435.
PA (UABR-) UAB RES FOUND.
PI Garver RI, Sorscher EJ.
PI WPI: 94-316537/39.
DR DNA construct for treating human carcinoma - includes a
PT cancer-therapeutic gene under the control of a promoter and a gp.
PT of enhancer sequences
PS Claim 1; Fig. 6; 54pp; English.
CC This enhancer element is part of a DNA construct used for treating
CC human carcinoma which contains a cancer therapeutic protein under
CC the control of a promoter and 3 enhancer sequences in a specific 5'-3'
CC order. This enhancer element is derived from the flanking region of
CC the human epithelial cell cyokeratin-8 gene.
SQ Sequence 32 BP; 7 A; 1 C; 8 G; 15 T;
```

```
Query Match 0.1%; Score 28; DB 1; Length 32;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
05168/c
ID 05168 standard; DNA: 40 BP.
AC 05168;
DE 21-JUL-1994 (first entry)
DE Sequence of primer for PCR amplification of HIV-LP Pt.1 isolate
KW Human immunodeficiency virus; HIV-LP; PCR primer; ss.
OS Synthetic.
PN W09400562-A.
PD 06-JAN-1994.
PR 23-JUN-1993; U06162.
PR 24-JUN-1992; US-903421.
PA (CORR ) CORRELL RES FOUND INC.
PA (ORNT ) UNIV NEW YORK MT SINAI.
PA SCHOOL MEDICINE.
PI Gelman IH, Laurence JC;
PI WPI: 94-026200/03.
PT HIV-LP useful in vaccine formulations - is novel HIV virus
PT distinct from HIV-1 and or HIV-2 viruses
PS Example; page 7; 75pp; English.
CC HIV-LP is a new variant of the HIV family. A cDNA first strand was
CC synthesised from Pt. 1 pellet using MLV RT. The product was
CC converted into dsDNA and this cDNA was amplified by PCR using
CC primers 05167 and 05168.
SQ Sequence 40 BP; 3 A; 2 C; 3 G; 32 T;
```

```
RESULT 11
ID 05168/c
AC 05168;
DE 21-JUL-1994 (first entry)
DE Sequence of primer for PCR amplification of HIV-LP Pt.1 isolate
KW Human immunodeficiency virus; HIV-LP; PCR primer; ss.
OS Synthetic.
PN W09400562-A.
PD 06-JAN-1994.
PR 23-JUN-1993; U06162.
PR 24-JUN-1992; US-903421.
PA (CORR ) CORRELL RES FOUND INC.
PA (ORNT ) UNIV NEW YORK MT SINAI.
PA SCHOOL MEDICINE.
PI Gelman IH, Laurence JC;
PI WPI: 94-026200/03.
PT HIV-LP useful in vaccine formulations - is novel HIV virus
PT distinct from HIV-1 and or HIV-2 viruses
PS Example; page 7; 75pp; English.
CC HIV-LP is a new variant of the HIV family. A cDNA first strand was
CC synthesised from Pt. 1 pellet using MLV RT. The product was
CC converted into dsDNA and this cDNA was amplified by PCR using
CC primers 05167 and 05168.
SQ Sequence 40 BP; 3 A; 2 C; 3 G; 32 T;
```

```
Query Match 0.1%; Score 28; DB 1; Length 40;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
4696 AAAAAAAAAAAAAAAAAAAGAT 4723
DB 34 AAAAAAAAAAAAAAAAAAAGAT 7
```

```
RESULT 12
ID T04081/c
AC T04081 standard; DNA: 48 BP.
DE 16-MAR-1996 (first entry)
DE Trypsin inhibitor protein cDNA antisense primer-1.
KW Trypsin; inhibitor; human 198G cells; pancreatitis; shock; DIC;
KW multiple organ failure; disseminated intravascular coagulation; ss.
OS Synthetic.
PN J07242700-A.
PD 19-SEP-1995.
PR 04-MAR-1994; 059906.
PR 04-MAR-1994; JP-059906.
PA (MOCH ) MOCHIDA PHARM CO LTD.
PA WPI: 95-355285/46.
DR Polypeptide having trypsin inhibitory activity - for the treatment
PT of pancreatitis, shock, multi-organ failure, etc.
PS Example 7; Page 15; 25pp; Japanese.
CC A polypeptide having trypsin inhibitory activity, mol.wt. 22-28 kD
CC or 17-23 kD (as determined by SDS-PAGE under reducing or non-
CC reducing conditions, respectively) and an N-terminal sequence as in
CC R/913 is claimed. The polypeptide was isolated from human 198G
CC cells and is useful for treating pancreatitis, shock, multiple
CC organ failure and disseminated intravascular coagulation. The
CC present sequence is that of an antisense primer used for cloning
CC 198G cDNA coding for the full-length polypeptide.
SQ Sequence 48 BP; 2 A; 1 C; 1 G; 44 T;
```


Query Match 0.1%; Score 28; DB 1; Length 48;
Best Local Similarity 100.0%; Pred. No. 43;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4696 AAAAAAAAAAAAAAAAAAGAAAT 4723
DB 32 AAAAAAAAAAAAAAAAAAGAAAT 5

RESULT 13
V02146 standard; CDNA; 49 BP.
AC V02146;
DE 12-MAY-1998 (first entry)
KW Human secreted protein AK533 3' portion including the polyA tail.
KW Human secreted protein; ATCC 98026; cytokine; immunomodulation;
KW cell proliferation; differentiation; regulation; ds.
OS Homo sapiens.
PN WO9739123-A2.
PD 23-OCT-1997.
PF 14-APR-1997; U06139.
PR 18-APR-1996; US-634325.
PA (GENE) GENETICS INST INC.
PI Jacobs K, Lavallee ER, McCoy JM, Merberg D, Racie LA, Spaulding V, WPI; 97-526459/48.
DR New secreted proteins encoded clones present in ATCC 98026 - possibly having cytokine, cell proliferation/differentiation regulating, immunomodulating and many other activities
PS Disclosure; Page 86; 139pp; English.
CC The present sequence encodes a portion of a novel human secreted protein deposited under accession number ATCC 98026. The secreted protein can be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions and as nutritional supplements. It may also have a very wide range of biological activities although no evidence for any is provided in the specification. Typical of these are induction of other cytokines; immunostimulating/immunosuppressant activities (e.g. for treating human immunodeficiency virus infection, cancer, autoimmune diseases and allergy); regulation of hematopoiesis (e.g. for treating anaemia or as adjunct to chemotherapy); stimulation of growth of bone, cartilage, tendons, ligaments and/or nerves (e.g. for treating wounds, periodontal disease, neurological diseases stroke, fibrosis); inhibition or stimulation of follicle stimulating hormone (for control of fertility); chemotactic and chemokinetic activities (e.g. for treating infections, tumours); haemostatic or thrombolytic activity (e.g. for treating haemophilia, cardiac infarction etc.); anti-inflammatory activity (e.g. for treating septic shock, Crohn's disease); as antimicrobials; for treating postnatal or other hyperproliferative disease; for regulation of metabolism, behaviour, and many others. Also contemplated is the use of the corresponding nucleic acid in gene therapy procedures.
Sequence 49 BP; 40 A; 2 C; 6 G; 1 T;

Query Match 0.1%; Score 28; DB 1; Length 49;
Best Local Similarity 100.0%; Pred. No. 43;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 26678 GAAAAAAAAAAAAAAAAAAAAA 26705
DB 3 GAAAAAAAAAAAAAAAAAAAAA 30

RESULT 14
T88080 standard; CDNA; 49 BP.
AC T88080;
DE 14-MAY-1998 (first entry)
PT 3' portion of cDNA clone encoding secreted protein AK533.

KW Human; secreted protein; research; treatment; AK533;
KW 3' portion; ds.
OS Homo sapiens.
PN WO9739122-A2.
PD 23-OCT-1997.
PF 11-APR-1997; U06042.
PR 12-APR-1996; US-631184.
PA (MORO) MORO PHARM INC.
PI Thecharides TC; WPI; 97-526459/48.
DR Human and murine secreted proteins - useful to research or treat diseases or disorders related to their function
PS Disclosure; Page 86; 140pp; English.
CC The present sequence is the 3' portion of a cDNA clone encoding a human secreted protein, which may have nutritional uses, or cytokine and cell proliferation/differentiation, immune stimulating or suppressing, haematopoiesis regulating, tissue growth, activin/inhibin, chemotactic/chemokinetic, haemostatic and thrombolytic, receptor/ligand, anti-inflammatory or tumour inhibition activities. It can also be used to research or treat diseases/disorders related to its function.
CC The partial cDNA clone AP162 was 1st isolated from a human adult placenta cDNA library. The partial cDNA clones AM931, AM610, AM340, AM282, AK647, AK583, AK533 and AK296 were 1st isolated from a human foetal kidney cDNA library. The partial cDNA clones H617 and B89 or Th2) cDNA library. The partial cDNA clone AW191 was 1st isolated from a human ovary (PA-1 teratocarcinoma) cDNA library. The partial cDNA clones AT211, AT205 and AT319 were 1st isolated from a human lymphocyte and dendritic cell cDNA library. The partial cDNA clones AS34 and AS32 were 1st isolated from a human foetal brain cDNA library. The partial cDNA clone AR260 was 1st isolated from a human adult retina cDNA library. The partial cDNA clones K640 and K39 were 1st isolated from a murine bone marrow (stromal cell line FCW-4) cDNA library.
Sequence 49 BP; 40 A; 2 C; 6 G; 1 T;

Query Match 0.1%; Score 28; DB 1; Length 49;
Best Local Similarity 100.0%; Pred. No. 43;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 26678 GAAAAAAAAAAAAAAAAAAAAA 26705
DB 3 GAAAAAAAAAAAAAAAAAAAAA 30

RESULT 15
V54636/c
ID V54636 standard; DNA; 53 BP.
AC V54636;
DE 18-NOV-1998 (first entry)
PT DNA detector sequence used in a telomeric activity assay.
KW Telomerase; telomeric repeat sequence; carcinogen; senescence; cancer; ss.
OS Synthetic.
FH Key
FT misc_binding
FT Location/Qualifiers
FT 31..54
FT /tag= a
FT /note= "Forms a double-stranded region with bases 38-60 of V54639".
PN WO9837241-A1.
PD 27-AUG-1998.
PF 23-FEB-1998; U03725.
PR 18-JUN-1997; US-050109.
PR 24-FEB-1997; US-038798.
PA (TMTE-) TM TECHNOLOGIES INC.
PI Benight AS, Faldasz BD, Lane MJ; WPI; 98-467588/40.
DR Detecting telomerase activity - without the need for polymerase chain reaction amplification, can be adapted to assay for telomerase inhibitors
PT Example 1; Fig 1C; 9pp; English.

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 30, 2000, 23:51:03 ; Search time 339.96 Seconds
(without alignments)
10565.402 Million cell updates/sec

Title: US-08-852-495c-1_COPY_140000_170000

Perfect score: 30001
Sequence: 1 CCCCTCCATCCCATATGCAC.....TACATATTTATTTATGTA 30001

Scoring table:
OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 214294 segs, 59861574 residues

Word size : 0

Total number of hits satisfying chosen parameters: 349160

Minimum DB seq length: 8
Maximum DB seq length: 105

Post-processing: Listing first 45 summaries

Database : Issued_Patents_NA.*
1: /cgn2_6/p10data/1/lna/5A_COMB.seq.*
2: /cgn2_6/p10data/1/lna/5B_COMB.seq.*
3: /cgn2_6/p10data/1/lna/5C_COMB.seq.*
4: /cgn2_6/p10data/1/lna/5D_COMB.seq.*
5: /cgn2_6/p10data/1/lna/6_COMB.seq.*
6: /cgn2_6/p10data/1/lna/PCRUS_COMB.seq.*
7: /cgn2_6/p10data/1/lna/backfile1.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	38	0.1	60	3	US-08-454-557C-57
2	38	0.1	60	4	US-08-340-426D-57
3	38	0.1	60	4	US-08-450-673C-57
4	38	0.1	60	6	PCT-US95-17111A-57
5	38	0.1	76	3	US-08-454-557C-69
6	38	0.1	76	4	US-08-340-426D-69
7	38	0.1	76	4	US-08-450-673C-69
8	38	0.1	76	6	PCT-US95-17111A-69
9	29	0.1	30	3	US-08-454-557C-6
10	29	0.1	30	4	US-08-340-426D-6
11	29	0.1	30	4	US-08-450-673C-6
12	29	0.1	30	6	PCT-US95-17111A-6
13	29	0.1	55	4	US-08-778-494B-67
14	29	0.1	84	3	US-08-778-494B-109
15	29	0.1	84	3	US-08-454-557C-91
16	29	0.1	84	4	US-08-340-426D-91
17	29	0.1	84	4	US-08-450-673C-91
18	29	0.1	84	6	PCT-US95-17111A-91
19	28	0.1	50	1	US-08-233-609-5
20	28	0.1	50	1	US-08-444-083-5
21	28	0.1	50	1	US-08-286-304-5
22	28	0.1	50	1	US-08-442-745-5
23	28	0.1	50	1	US-08-443-129-5
24	28	0.1	50	1	US-08-443-952-5
25	28	0.1	50	2	US-08-443-130-5
26	28	0.1	50	6	PCT-US95-04467-5

27	28	0.1	60	3	US-08-454-557C-58	Sequence 58, Appl
28	28	0.1	60	3	US-08-454-557C-58	Sequence 58, Appl
29	28	0.1	60	4	US-08-340-426D-58	Sequence 58, Appl
30	28	0.1	60	4	US-08-340-426D-58	Sequence 58, Appl
31	28	0.1	60	4	US-08-450-673C-58	Sequence 58, Appl
32	28	0.1	60	4	US-08-450-673C-58	Sequence 58, Appl
33	28	0.1	60	6	PCT-US95-17111A-58	Sequence 58, Appl
34	28	0.1	60	6	PCT-US95-17111A-58	Sequence 58, Appl
35	28	0.1	78	3	US-08-454-557C-70	Sequence 70, Appl
36	28	0.1	78	3	US-08-454-557C-70	Sequence 70, Appl
37	28	0.1	78	4	US-08-340-426D-70	Sequence 70, Appl
38	28	0.1	78	4	US-08-340-426D-70	Sequence 70, Appl
39	28	0.1	78	4	US-08-450-673C-70	Sequence 70, Appl
40	28	0.1	78	4	US-08-450-673C-70	Sequence 70, Appl
41	28	0.1	78	6	PCT-US95-17111A-70	Sequence 70, Appl
42	28	0.1	78	6	PCT-US95-17111A-70	Sequence 70, Appl
43	27	0.1	33	7	5478746-1	Patent No. 5478746
44	27	0.1	42	1	US-07-875-167-2	Sequence 2, Appl1
45	27	0.1	42	1	US-08-287-164-2	Sequence 2, Appl1

ALIGNMENTS

RESULT 1
US-08-454-557C-57
Sequence 57, Application US/08454557C
Patent No. 5830670

GENERAL INFORMATION:
APPLICANT: de la Monte, Suzanne

TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection

TITLE OF INVENTION: of Alzheimer's Disease

NUMBER OF SEQUENCES: 121

CORRESPONDENCE ADDRESS:

ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.

STREET: 1100 New York Avenue, Suite 600

CITY: Washington

STATE: D.C.

COUNTRY: U.S.A.

ZIP: 20005-3934

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/454, 557C

FILING DATE: 30-MAY-1995

CLASSIFICATION: 514

AUTHOR/AGENT INFORMATION:
NAME: Ludwig, Steven R.

REGISTRATION NUMBER: 36, 203

REFERENCE/DOCKET NUMBER: 0609.3840003

TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600

TELEFAX: (202) 371-2540

INFORMATION FOR SEQ ID NO: 57:

SEQUENCE CHARACTERISTICS:

LENGTH: 60 base pairs

TYPE: nucleic acid

STRANDEDNESS: both

TOPOLOGY: both

US-08-454-557C-57

Query Match 0.1%; Score 38; DB 3; Length 60;
Best Local Similarity 100.0%; Pred. No. 0.0042;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9496 GCTAATTTTGTATTTTGTAGACAGACGGTTTCACC 9533
|||||
Db 11 GCTAATTTTGTATTTTGTAGACAGACGGTTTCACC 48

MEDIUM TYPE: Floppy disk

TOPOLOGY: both

PCT-US95-17111A-57

Query Match 0.1%; Score 38; DB 6; Length 60;
Best Local Similarity 100.0%; Pred. No. 0.0042;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9496 GCTAATTTTGTATTAGTAGACAGAGGTTTCACC 9533
|||||
Db 11 GCTAATTTTGTATTAGTAGACAGAGGTTTCACC 48

RESULT 5

US-08-454-557C-69
; Sequence 69, Application US/08454557C
; Patent No. 5830670
; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; TITLE OF INVENTION: Neutral Thread Protein Gene Expression and Detection
; TITLE OF INVENTION: of Alzheimer's Disease
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/454,557C
; FILING DATE: 30-MAY-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Ludwig, Steven R.
; REGISTRATION NUMBER: 36,203
; REFERENCE/DOCKET NUMBER: 0609.3840003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 69:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 76 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
US-08-454-557C-69

Query Match 0.1%; Score 38; DB 3; Length 76;
Best Local Similarity 100.0%; Pred. No. 0.004;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9496 GCTAATTTTGTATTAGTAGACAGAGGTTTCACC 9533
|||||
Db 11 GCTAATTTTGTATTAGTAGACAGAGGTTTCACC 48

RESULT 6

US-08-340-426D-69
; Sequence 69, Application US/08340426D
; Patent No. 5948634
; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; TITLE OF INVENTION: Neutral Thread Protein Gene Expression and Detection
; TITLE OF INVENTION: of Alzheimer's Disease
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESS:

ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934

COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/340,426D
; FILING DATE: 14-NOV-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Ludwig, Steven R.
; REGISTRATION NUMBER: 36,203
; REFERENCE/DOCKET NUMBER: 0609.3840002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 69:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 76 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
US-08-340-426D-69

Query Match 0.1%; Score 38; DB 4; Length 76;
Best Local Similarity 100.0%; Pred. No. 0.004;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9496 GCTAATTTTGTATTAGTAGACAGAGGTTTCACC 9533
|||||
Db 11 GCTAATTTTGTATTAGTAGACAGAGGTTTCACC 48

RESULT 7

US-08-450-673C-69
; Sequence 69, Application US/08450673C
; Patent No. 5948888
; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; TITLE OF INVENTION: Neutral Thread Protein Gene Expression and Detection
; TITLE OF INVENTION: of Alzheimer's Disease
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/450,673C
; FILING DATE: 30-MAY-1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Ludwig, Steven R.
; REGISTRATION NUMBER: 36,203
; REFERENCE/DOCKET NUMBER: 0609.3840004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540

FILING DATE: 14-NOV-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609.3840002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 30 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
US-08-340-426D-6

Query Match 0.1%; Score 29; DB 4; Length 30;
Best Local Similarity 100.0%; Pred. No. 2.5;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28388 TTCAGCGATTCTCTGCTCAGCTCC 28416
DB 30 TTCAGCGATTCTCTGCTCAGCTCC 2

RESULT 11
US-08-450-673C-6/C
Sequence 6, Application US/08450673C
Patent No. 594888
GENERAL INFORMATION:
APPLICANT: de la Monte, Suzanne
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609.3840002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 30 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
US-08-450-673C-6

Query Match 0.1%; Score 29; DB 4; Length 30;
Best Local Similarity 100.0%; Pred. No. 2.5;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28388 TTCAGCGATTCTCTGCTCAGCTCC 28416
DB 30 TTCAGCGATTCTCTGCTCAGCTCC 2

RESULT 12
PCT-US95-17111A-6/C
Sequence 6, Application PC/TUS9517111A
GENERAL INFORMATION:
APPLICANT: de la Monte, Suzanne
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609.3840002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 30 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
PCT-US95-17111A-6

Query Match 0.1%; Score 29; DB 6; Length 30;
Best Local Similarity 100.0%; Pred. No. 2.5;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28388 TTCAGCGATTCTCTGCTCAGCTCC 28416
DB 30 TTCAGCGATTCTCTGCTCAGCTCC 2

RESULT 13
US-08-778-494B-67/C
Sequence 67, Application US/08778494B
Patent No. 5962272
GENERAL INFORMATION:
APPLICANT: Chenchik, Alex
ATTORNEY/AGENT INFORMATION:
NAME: Zhu, York
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609.3840002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 30 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
US-08-778-494B-67/C

CITY: Gainesville
STATE: Florida
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/778,494B
FILING DATE: 03-JAN-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/582,562
FILING DATE: 03-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Pace, Doran R.
REGISTRATION NUMBER: 38,261
REFERENCE/DOCKET NUMBER: CL-7C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (352) 375-8100
TELEFAX: (352) 372-5800
INFORMATION FOR SEQ ID NO: 67:
SEQUENCE CHARACTERISTICS:
LENGTH: 55 bases
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (synthetic)
US-08-778-494B-67

Query Match 0.1%; Score 29; DB 4; Length 55;
Best Local Similarity 100.0%; Pred. No. 2.1;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 26680 AAAAAAAAAAAAAAAAAAGTA 26708
DB 51 AAAAAAAAAAAAAAAAAAGTA 23

RESULT 14
US-08-778-494B-109/C
Sequence 109; Application US/08/78494B
Patent No. 5962272
GENERAL INFORMATION:
APPLICANT: Chenchik, Alex
APPLICANT: Zhu, Yohk
APPLICANT: Diachenko, Luda
APPLICANT: Siebert, Paul
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR FULL-LENGTH CDNA
TITLE OF INVENTION: CLONING
NUMBER OF SEQUENCES: 114
CORRESPONDENCE ADDRESS:
ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: Florida
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/778,494B
FILING DATE: 03-JAN-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/582,562
FILING DATE: 03-JAN-1996

ATTORNEY/AGENT INFORMATION:
NAME: Pace, Doran R.
REGISTRATION NUMBER: 38,261
REFERENCE/DOCKET NUMBER: CL-7C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (352) 375-8100
TELEFAX: (352) 372-5800
INFORMATION FOR SEQ ID NO: 109:
SEQUENCE CHARACTERISTICS:
LENGTH: 58 bases
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (synthetic)
US-08-778-494B-109

Query Match 0.1%; Score 29; DB 4; Length 58;
Best Local Similarity 100.0%; Pred. No. 2.1;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 26680 AAAAAAAAAAAAAAAAAAGTA 26708
DB 52 AAAAAAAAAAAAAAAAAAGTA 24

RESULT 15
US-08-454-557C-91/C
Sequence 91; Application US/0845457C
Patent No. 5830670
GENERAL INFORMATION:
APPLICANT: de la Monte, Suzanne
APPLICANT: Mandis, Jack R.
TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection
TITLE OF INVENTION: of Alzheimer's Disease
NUMBER OF SEQUENCES: 121
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3924
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/454,557C
FILING DATE: 30-MAY-1995
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609.3840003
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 91:
SEQUENCE CHARACTERISTICS:
LENGTH: 84 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
US-08-454-557C-91

QY 24038 TGTATCCAGCATTGTGGAGCTGAG 24066

Mon Apr 3 08:24:08 2000

us-08-852-495c-1_copy_140000_170000.rni

Db 77 TGTATCCAGCAGCTTTGGAGGCTGAGG 49

Search completed: March 31, 2000, 10:04:17
Job time: 195319 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 30, 2000, 19:57:22 ; Search time 8179.71 Seconds
(without alignments)
13848.131 Million cell updates/sec

Title: us-08-852-495c-1_COPY_140000_170000
Perfect score: 30001
Sequence: 1 CCCGCCAATCCCATATGCAC.....TACATATTATATTATGTA 30001

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 4538634 seqs, 188781982 residues

Word size : 0

Total number of hits satisfying chosen parameters: 130662

Minimum DB seq length: 8
Maximum DB seq length: 105

Post-processing: Listing first 45 summaries

Database : EST.*
1: em_est1:*
2: em_est2:*
3: em_est3:*
4: em_est4:*
5: em_est5:*
6: em_est6:*
7: em_est7:*
8: em_est8:*
9: em_est9:*
10: em_est10:*
11: em_est11:*
12: em_est12:*
13: em_est13:*
14: em_est14:*
15: em_est15:*
16: em_est16:*
17: em_est17:*
18: em_est18:*
19: em_est19:*
20: gb_est1:*
21: gb_est2:*
22: gb_est3:*
23: gb_est4:*
24: gb_est5:*
25: gb_est6:*
26: gb_est7:*
27: gb_est8:*
28: gb_est9:*
29: gb_est10:*
30: gb_est11:*
31: gb_est12:*
32: gb_est13:*
33: gb_est14:*
34: gb_est15:*
35: gb_est16:*
36: gb_est17:*
37: gb_est18:*
38: gb_est19:*
39: gb_est20:*
40: gb_est21:*
41: gb_est22:*
42: gb_est23:*
43: gb_est24:*

44: gb_est25:*
45: gb_est26:*
46: gb_est27:*
47: gb_est28:*
48: gb_est29:*
49: gb_est30:*
50: gb_est31:*
51: gb_est32:*
52: em_est20:*
53: em_est21:*
54: em_est22:*
55: em_est23:*
56: em_est24:*
57: em_est25:*
58: em_est26:*
59: gb_est33:*
60: gb_est34:*
61: gb_est35:*
62: gb_est36:*
63: gb_est37:*
64: gb_est38:*
65: em_est27:*
66: em_est28:*
67: em_est29:*
68: em_est30:*
69: gb_est39:*
70: gb_est40:*
71: gb_est41:*
72: gb_est42:*
73: gb_est43:*
74: gb_est44:*
75: em_est31:*
76: em_est32:*
77: em_est33:*
78: em_est34:*
79: gb_gss1:*
80: gb_gss2:*
81: gb_gss3:*
82: gb_gss4:*
83: em_gss1:*
84: em_gss2:*
85: em_gss3:*
86: em_gss4:*
87: gb_gss5:*
88: gb_gss6:*
89: gb_gss7:*
90: gb_gss8:*
91: gb_gss9:*
92: em_gss5:*
93: em_gss6:*
94: em_gss7:*
95: em_gss8:*
96: em_gss9:*
97: em_gss10:*
98: em_gss11:*
99: gb_gss10:*
100: gb_gss11:*
101: em_gss12:*
102: gb_gss12:*
103: gb_gss13:*
104: gb_gss14:*
105: gb_gss15:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	45	0.1	97	25	N47846 yw95g08.r1

```

c 2 45 0.1 102 81 B62983
c 3 43 0.1 102 81 B48088
c 4 41 0.1 61 29 AA873656
c 5 41 0.1 95 20 T52775
c 6 40 0.1 72 37 AA715942
c 7 40 0.1 105 21 T63772
c 8 39 0.1 97 39 AA837701
c 9 39 0.1 102 81 B80126
c 10 38 0.1 104 27 C05051
c 11 37 0.1 80 50 F24490
c 12 37 0.1 96 28 AA078527
c 13 37 0.1 100 30 AA252633
c 14 37 0.1 71 25 N80946
c 15 36 0.1 101 25 N80746
c 16 36 0.1 87 20 D16500
c 17 36 0.1 95 40 AA923035
c 18 36 0.1 102 25 N83396
c 19 35 0.1 92 37 AA729064
c 20 35 0.1 95 35 AA578401
c 21 35 0.1 102 23 H39052
c 22 35 0.1 105 61 A1832832
c 23 35 0.1 105 99 A0196366
c 24 34 0.1 83 40 AA988425
c 25 34 0.1 87 20 D20989
c 26 34 0.1 87 21 T62174
c 27 34 0.1 37 22 R70733
c 28 33 0.1 77 62 A1914818
c 29 33 0.1 86 38 AA745546
c 30 33 0.1 100 34 AA515337
c 31 33 0.1 101 35 AA583697
c 32 33 0.1 102 81 B62983
c 33 33 0.1 104 39 AA831816
c 34 33 0.1 105 51 A1752407
c 35 32 0.1 80 31 AA289455
c 36 32 0.1 86 38 AA745546
c 37 32 0.1 102 36 AA654896
c 38 32 0.1 104 38 AA807731
c 39 32 0.1 105 81 B74378
c 40 31 0.1 53 43 A1206968
c 41 31 0.1 53 43 A1242004
c 42 31 0.1 53 44 A1249836
c 43 31 0.1 53 44 A1252536
c 44 31 0.1 54 44 AA457759
c 45 31 0.1 55 23 H51371

```

ALIGNMENTS

```

RESULT 1
LOCUS N47846 97 bp mRNA EST 14-FEB-1996
DEFINITION y959g08.r1 Soares,placenta,8to9weeks_2NBHP8to9w Homo sapiens cDNA
clone IMAGE:260030.5' similar to contains Alu repetitive element;
mRNA sequence.
N47846
N47846.1 GI:1189012

```

```

ACCESSION N47846
VERSION N47846
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

```

```

REFERENCE 1 (bases 1 to 97)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS Hallier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
Holman,M., Hultman,M., Kueba,T., Le,M., Lennon,G., Marra,M.,
Parsons,J., Rifkin,J., Rohlfing,T., Soares,M., Tan,F.,
Trevaskis,E., Waterston,R., Williamson,A., Woldmann,P. and
Wilson,R.

```

```

TITLE The Mashu-Merck EST Project
JOURNAL Unpublished (1995)
COMMENT On May 9, 1995 this sequence version replaced gi:802689.
Contact: Wilson RK
Washington University School of Medicine

```

```

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.edu
This clone is available royalty-free through LNC; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Putative full length read
Seq primer: T7
High quality sequence stop: 129.
Location/Qualifiers
1..97
/organism="Homo sapiens"
/db_xref="GDB:3889736"
/db_xref="taxon:9606"
/clone="IMAGE:260030"
/cdate="1995"
/clone_1lb="Soares,placenta,8to9weeks_2NBHP8to9w"
/clone_2lb="Soares,placenta,8to9weeks_2NBHP8to9w"
/clone_3lb="Soares,placenta,8to9weeks_2NBHP8to9w"
/clone_4lb="Soares,placenta,8to9weeks_2NBHP8to9w"
/clone_5lb="Soares,placenta,8to9weeks_2NBHP8to9w"
/clone_6lb="Soares,placenta,8to9weeks_2NBHP8to9w"
/clone_7lb="Soares,placenta,8to9weeks_2NBHP8to9w"
/clone_8lb="Soares,placenta,8to9weeks_2NBHP8to9w"
/clone_9lb="Soares,placenta,8to9weeks_2NBHP8to9w"
/clone_10lb="Soares,placenta,8to9weeks_2NBHP8to9w"
/clone_11lb="Soares,placenta,8to9weeks_2NBHP8to9w"
/clone_12lb="Soares,placenta,8to9weeks_2NBHP8to9w"
/clone_13lb="Soares,placenta,8to9weeks_2NBHP8to9w"
/clone_14lb="Soares,placenta,8to9weeks_2NBHP8to9w"
/clone_15lb="Soares,placenta,8to9weeks_2NBHP8to9w"
/clone_16lb="Soares,placenta,8to9weeks_2NBHP8to9w"
/clone_17lb="Soares,placenta,8to9weeks_2NBHP8to9w"
/clone_18lb="Soares,placenta,8to9weeks_2NBHP8to9w"
/clone_19lb="Soares,placenta,8to9weeks_2NBHP8to9w"
/clone_20lb="Soares,placenta,8to9weeks_2NBHP8to9w"
/clone_21lb="Soares,placenta,8to9weeks_2NBHP8to9w"
/clone_22lb="Soares,placenta,8to9weeks_2NBHP8to9w"
/clone_23lb="Soares,placenta,8to9weeks_2NBHP8to9w"
/clone_24lb="Soares,placenta,8to9weeks_2NBHP8to9w"
/clone_25lb="Soares,placenta,8to9weeks_2NBHP8to9w"
/clone_26lb="Soares,placenta,8to9weeks_2NBHP8to9w"
/clone_27lb="Soares,placenta,8to9weeks_2NBHP8to9w"
/clone_28lb="Soares,placenta,8to9weeks_2NBHP8to9w"
/clone_29lb="Soares,placenta,8to9weeks_2NBHP8to9w"
/clone_30lb="Soares,placenta,8to9weeks_2NBHP8to9w"
/clone_31lb="Soares,placenta,8to9weeks_2NBHP8to9w"
/clone_32lb="Soares,placenta,8to9weeks_2NBHP8to9w"
/clone_33lb="Soares,placenta,8to9weeks_2NBHP8to9w"
/clone_34lb="Soares,placenta,8to9weeks_2NBHP8to9w"
/clone_35lb="Soares,placenta,8to9weeks_2NBHP8to9w"
/clone_36lb="Soares,placenta,8to9weeks_2NBHP8to9w"
/clone_37lb="Soares,placenta,8to9weeks_2NBHP8to9w"
/clone_38lb="Soares,placenta,8to9weeks_2NBHP8to9w"
/clone_39lb="Soares,placenta,8to9weeks_2NBHP8to9w"
/clone_40lb="Soares,placenta,8to9weeks_2NBHP8to9w"
/clone_41lb="Soares,placenta,8to9weeks_2NBHP8to9w"
/clone_42lb="Soares,placenta,8to9weeks_2NBHP8to9w"
/clone_43lb="Soares,placenta,8to9weeks_2NBHP8to9w"
/clone_44lb="Soares,placenta,8to9weeks_2NBHP8to9w"
/clone_45lb="Soares,placenta,8to9weeks_2NBHP8to9w"

```

```

FEATURES
SOURCE
BASE COUNT 23 a 22 c 32 g 20 t
ORIGIN
Query Match 0.1%; Score 45; DB 25; Length 97;
Best Local Similarity 100.0%; Pred. No. 0.00096;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Oy 19136 CACCACTGCTCCAGCTGCGGCGAGGAGAGCTGTCTCA 19180
Db 53 CACCACTGCTCCAGCTGCGGCGAGGAGAGCTGTCTCA 97

```

```

RESULT 2
LOCUS B62983 102 bp DNA GSS 21-JUN-1998
DEFINITION CIT-HSP-669F15.TP CIT-HSP Homo sapiens genomic clone 669F15,
genomic survey sequence.
B62983
B62983.1 GI:2636892

```

```

ACCESSION B62983
VERSION B62983
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 102)
Kim,U.-J., Adams,M.D. and Simon,M.I.
Determination of clone end sequences of human Bacterial Artificial
Chromosomes
Unpublished (1997)
Contact: Ung-Jin Kim
California Institute of Technology
Division of Biology, MS 147-75, Pasadena, CA 91125, USA
Tel: 626 796 7066
Fax: 626 395 4901
Email: ung@caltech.edu
Clones are available from Research Genetics (info@resgen.com). BAC
and search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
Seq primer: SP6
Class: BAC ends.
Location/Qualifiers
1..102
/organism="Homo sapiens"

```

```

FEATURES
SOURCE

```

/db_xref="GDB:5491647"
/db_xref="taxon:9606"
/clone_1lb="CIT-HSP"
/sex="Male"
/cell_type="Sperm"
/note="Vector: pBelobAC1; site_1: HindIII; site_2: HindIII"

BASE COUNT 30 a 30 c 21 g 21 t
ORIGIN

Query Match 0.1%; Score 45; DB 81; Length 102;
Best Local Similarity 100.0%; Pred. No. 0.00093; Indels 0; Gaps 0;
Matches 45; Conservative 0; Mismatches 0;

Oy 28457 GCTAATTTTGTATTTAGTAGAGAGGGGTTTACCAGTTGG 28501
|||||
Db 64 GCTAATTTTGTATTTAGTAGAGAGGGGTTTACCAGTTGG 20

RESULT 3
B48088/c 102 bp DNA GSS 08-APR-1999
LOCUS RPCI11-4N6.TP RPCI-11 Homo sapiens genomic clone RPCI-11-4N6,
DEFINITION genomic survey sequence.
ACCESSION B48088
VERSION B48088.1 GI:2600325
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 102)
AUTHORS Adams,M.D., Rounsley,S.D., Field,C.E., Bass,S., Linher,K.,
Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., de Jong,P. and
Venter,J.C.
TITLE Use of BAC End Sequences for Sequence-Ready Map Building
JOURNAL unpublished (1997)
COMMENT Other_GSSs: RPCI11-4N6.TV
Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Clones are derived from the human BAC library RPCI-11. For BAC
library availability, please contact Pieter de Jong
(pieter@telong.med.buffalo.edu). Clones may be purchased from
BACRPC Resources (<http://bacrpc.med.buffalo.edu/ordering>) or from
Research Genetics (inforesgen.com). BAC end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
Seq primer: SP6
Class: BAC ends.
Location/Qualifiers
1. 102
/organism="Homo sapiens"
/db_xref="GDB:7501469"
/db_xref="taxon:9606"
/clone_1lb="RPCI-11-4N6"
/clone_1lb="RPCI-11"
/sex="Male"
/cell_type="Lymphocytes"
/note="Vector: pBAC3.6; site_1: EcoRI; site_2: EcoRI;
RPCI11 Human Male BAC library"

BASE COUNT 32 a 24 c 25 g 21 t
ORIGIN

Query Match 0.1%; Score 43; DB 81; Length 102;
Best Local Similarity 100.0%; Pred. No. 0.0038; Indels 0; Gaps 0;
Matches 43; Conservative 0; Mismatches 0;

Oy 9501 TTTTGTATTTTGTAGTAGAGAGGGGTTTACCAGTTGGCA 9543
|||||
Db 66 TTTTGTATTTTGTAGTAGAGAGGGGTTTACCAGTTGGCA 24

RESULT 4
AA873656/c 61 bp mRNA EST 07-APR-1998
LOCUS CE02607.s1 NCI-CGAP_Ov2 Homo sapiens cDNA clone IMAGE:1384764
DEFINITION similar to contains Alu repetitive element; mRNA sequence.
ACCESSION AA873656
VERSION AA873656.1 GI:2969778
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 61)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL unpublished (1997)
COMMENT On May 8, 1995 this sequence version replaced g1:801221.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Christopher A. Moskaluk, M.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: David B. Krizman, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.C.E. Consortium/LINt at:
www-bio.11n1.gov/bdrp/image/image.html
Insert Length: 553 Std Error: 0.00
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 51.
Location/Qualifiers
1. 61
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_1lb="IMAGE:1384764"
/clone_1lb="NCI-CGAP_Ov2"
/sex="female"
/tissue_type="ovary"
/lab_host="DH10B"
/note="Vector: PAMPI0; mRNA made from invasive ovarian
tumor. cDNA made by oligo-dT priming. Non-directionally
cloned. Size-selected on agarose gel. average insert size
600 bp. Reference: Krizman et al. (1996) Cancer Research
56:5380-5383."

BASE COUNT 9 a 15 c 20 g 17 t
ORIGIN

Query Match 0.1%; Score 41; DB 39; Length 61;
Best Local Similarity 100.0%; Pred. No. 0.021; Indels 0; Gaps 0;
Matches 41; Conservative 0; Mismatches 0;

Oy 19122 GTGACCGAGATCGACACCTGACCTCGGCGACA 19162
|||||
Db 61 GTGACCGAGATCGACACCTGACCTCGGCGACA 21

RESULT 5
T52775/c 95 bp mRNA EST 06-FEB-1995
LOCUS Y279008.r1 Stratagene ovary (#937217) Homo sapiens cDNA clone
DEFINITION IMAGE:67887 5' similar to contains Alu repetitive element; mRNA
sequence.
ACCESSION T52775

VERSION T52775.1 GI:654635
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
AUTHORS Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 95)
Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chapell, B.,
Chissoe, S., Dietrich, N., Dubuque, T., Favello, A., Gish, W.,
Hawkins, M., Hultman, M., Kucaba, T., Lacy, M., Le, N.,
Mardis, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L.,
Rohlfing, T., Scheinberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J.,
Trevaskis, E., Underwood, K., Wohlmann, P., Waterston, R., Wilson, R.
and Marra, M.
TITLE Generation and analysis of 280,000 human expressed sequence tags
JOURNAL Genome Res. 6 (9), 807-828 (1996)
MEDLINE 97044478
COMMENT Other ESTs: YA79408.s1
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.wustl.edu
Insert Size: 1013
High quality sequence stops: 82 Source: IMAGE Consortium, LNL This
clone is available royalty-free through LNL; contact the IMAGE
Consortium (info@image.llnl.gov) for further information.
Insert Length: 1013 Std Error: 0.00
Seq primer: M13R1
High quality sequence stop: 82.
Location/Qualifiers
1..95
/organism="Homo sapiens"
/db_xref="GDB:504144"
/db_xref="taxon:9606"
/clone_1db="IMAGE:67887"
/clone_1lb="Stratagene ovary (#937217)"
/sex="female"
/dev_stage="49 year old"
/lab_host="SOLR cells (kanamycin resistant)"
/note="Organ: ovary; Vector: Bluescript SK; Site: 1: EcoRI;
Site: 2: XhoI; Cloned unidirectionally. Primer: Oligo dt.
Total ovary tissue, normal, caucasian. Average insert
size: 0.8 kb; Uni-ZAP XR Vector: -5' adaptor sequence: 5'
GAATCGGACGAG 3' -3' adaptor sequence: 5'
CTCGAGTTTCTTTTCTTTT 3'."

BASE COUNT 19 a 32 c 20 g 24 t
ORIGIN

Query Match 0.1%; Score 41; DB 20; Length 95;
Best Local Similarity 100.0%; Pred. No. 0.016;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 732 AGGAGATCACTGACCTGGAGCAGAGGTTGACGTGAG 772
|||||
Db 84 AGGAGATCACTGACCTGGAGCAGAGGTTGACGTGAG 44

RESULT 6
LOCUS AA715942 72 bp mRNA EST 29-DEC-1997
DEFINITION nv77f06.r1 NCI-CCAP_Br4 Homo sapiens CDNA clone IMAGE:1235843
similar to contains Alu repetitive element; mRNA sequence.
ACCESSION AA715942
VERSION AA715942.1 GI:2728216
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homiidae; Homo.
COMMENT 1 (bases 1 to 72)

AUTHORS NCI-CCAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT On May 8, 1995 this sequence version replaced gi:800927.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Ilan Kirsch, M.D., Kristina A. Cole, M.D.,
Ph.D. student, Rodrigo F. Chuangu, M.D., Michael R. Emmert-Buck,
M.D., Ph.D.
CDNA Library Preparation: David B. Kitzman, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CCAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
www.bio.llnl.gov/dbp/image/image.html

Seq primer: -28m13 rev1 ET from Amersham
High quality sequence stop: 59.
Location/Qualifiers
1..72
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_1db="IMAGE:1235843"
/clone_1lb="NCI-CCAP_Br4"
/sex="female"
/tissue_type="normal ductal tissue"
/lab_host="DH10B"
/note="Organ: breast; Vector: pAMP10; mRNA made from
normal breast ductal tissue. cDNA made by oligo-dt
priming. Non-directionally cloned. Size-selected on
agarose gel, average insert size 600 bp. Reference:
Kitzman et al. (1996) Cancer Research 56:5380-5383."

BASE COUNT 15 a 22 c 20 g 15 t
ORIGIN

Query Match 0.1%; Score 40; DB 37; Length 72;
Best Local Similarity 100.0%; Pred. No. 0.039;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19138 CCAGTCACTCAGCCTGGGCGACAGAGGAGCTCTGTC 19177
|||||
Db 71 CCAGTCACTCAGCCTGGGCGACAGAGGAGCTCTGTC 32

RESULT 7
LOCUS T63772 105 bp mRNA EST 17-FEB-1995
DEFINITION YC21f01.s1 Stratagene lung (#937210) Homo sapiens CDNA clone
IMAGE:81337 3' similar to contains Alu repetitive element; mRNA
sequence.
ACCESSION T63772
VERSION T63772.1 GI:667637
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 105)
Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chapell, B.,
Chissoe, S., Dietrich, N., Dubuque, T., Favello, A., Gish, W.,
Hawkins, M., Hultman, M., Kucaba, T., Lacy, M., Le, N.,
Mardis, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L.,
Rohlfing, T., Scheinberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J.,
Trevaskis, E., Underwood, K., Wohlmann, P., Waterston, R., Wilson, R.
and Marra, M.
TITLE Generation and analysis of 280,000 human expressed sequence tags
JOURNAL Genome Res. 6 (9), 807-828 (1996)
MEDLINE 97044478
COMMENT Contact: Wilson RK
Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: estewatson.wustl.edu

High quality sequence stops: 69 Source: IMAGE Consortium, LNL. This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.lnl.gov) for further information.
Insert Length: 513 Std Error: 0.00
Seq primer: -21m13
High quality sequence stop: 69.

FEATURES

source

Location/Qualifiers

1. 105
/organism="Homo sapiens"
/db_xref="GDB:484954"
/db_xref="taxon:9606"
/clone="IMAGE:81337"
/clone_1lb="Stratagene lung (#937210)"
/sex="male"
/dev_stage="72 years"
/lab_host="SOLR cells (kanamycin resistant)"
/note="Organ: lung; Vector: paluescript SK; Site_1:
ECORI; Site_2: XhoI; Cloned unidirectionally. Primer:
Oligo dt. normal lung. Average insert size: 1.0 kb;
uni-zap xr vector: -5' adaptor sequence: 5' GAATTCGCGACGAG
3' -3' adaptor sequence: 5' CTCGAGTTTCTTTTCTTTT 3'"

BASE COUNT

17 a 34 c 29 g 25 t

ORIGIN

Query Match 0.1%; Score 40; DB 21; Length 105;
Best Local Similarity 100.0%; Pred. No. 0.031;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 28372 ACCTCCGCTCCCGGTTCAAGCATCTCTCTGCTCAGC 28411

Db 36 ACCTCCGCTCCCGGTTCAAGCATCTCTCTGCTCAGC 75

RESULT 8
AA837701/c 97 bp mRNA EST 07-APR-1998
LOCUS 0606c02.g1 NCI_CGAP_Ov2 Homo sapiens cDNA clone IMAGE:1385090
DEFINITION similar to contains element PTR5 repetitive element ;, mRNA
Sequence.

ACCESSION AA837701 GI:2912900
VERSION AA837701.1
KEYWORDS EST.
SOURCE human.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE

1 (bases 1 to 97)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index

JOURNAL

On Jan 17, 1998 this sequence version replaced g1:1900439.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550

COMMENT

Email: Robert.Strausberg@nih.gov
Tissue Procurement: Christopher A. Moskaluk, M.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: David B. Kitzman, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/INL at:
www-bio.lnl.gov/bdrip/image/image.html

Insert Length: 451 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 50.

FEATURES

source

Location/Qualifiers

1. 97
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1385090"
/clone_1lb="NCI_CGAP_Ov2"
/sex="female"
/tissue_type="ovary"
/lab_host="DH10B"
/note="Vector: pAMP10; mRNA made from invasive ovarian
tumor, cDNA made by oligo-dt priming. Non-directionally
cloned. Size-selected on agarose gel, average insert size
600 bp. Reference: Kitzman et al. (1996) Cancer Research
56:5380-5383."

BASE COUNT

13 a 24 c 29 g 31 t

ORIGIN

Query Match 0.1%; Score 39; DB 39; Length 97;
Best Local Similarity 100.0%; Pred. No. 0.066;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 734 GAGATCATTGAACCTGGGAGGAGGAGGTTGCACTGAG 772

Db 97 GAGATCATTGAACCTGGGAGGAGGAGGTTGCACTGAG 59

RESULT 9

B80126 102 bp DNA GSS 24-OCT-1998
LOCUS CIT-HSP-2042J24.TR CIT-HSP Homo sapiens genomic clone 2042J24,
DEFINITION genomic survey sequence.

ACCESSION B80126
VERSION B80126.1 GI:2867149
KEYWORDS GSS.
SOURCE human.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE

1 (bases 1 to 102)
Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
Simon,M. and Venter,J.C.
Use of a random BAC End Sequence Database for Sequence-Ready Map
Building (1998)

JOURNAL

Unpublished (1998)
Other_GSSs: CIT-HSP-2042J24.TR
Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208

COMMENT

Email: mdadams@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
Seq primer: M13 Reverse
Class: BAC ends.

FEATURES

source

Location/Qualifiers

1. 102
/organism="Homo sapiens"
/db_xref="GDB:705355"
/db_xref="taxon:9606"
/clone="2042J24"
/clone_1lb="CIT-HSP"
/sex="male"
/cell_type="Sperm"
/note="Vector: pBelobAC1; Site_1: HindIII; Site_2:
HindIII"

BASE COUNT 20 a 32 c 19 g 31 t

ORIGIN

Query Match 0.1%; Score 39; DB 81; Length 102;
 Best Local Similarity 100.0%; Pred. No. 0.065;
 Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 24024 CAGTGGCTCAGCGCTGTATCCAGACATTGGAGGCT 24062

Db 64 CAGTGGCTCAGCGCTGTATCCAGACATTGGAGGCT 102

RESULT 10
 C05051/c 104 bp mRNA EST 30-JUL-1996
 LOCUS C05051 Human heart cDNA (Ynakamura) Homo sapiens cDNA clone
 DEFINITION 3NHC4544, mRNA sequence.

ACCESSION C05051
 VERSION C05051.1 GI:1468302
 KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Homiidae; Homo.
 1 (bases 1 to 104)
 Tanaka,T., Ogihara,A., Uchiyama,I., Takagi,T., Yazaki,Y. and
 Nakamura,Y.

TITLE Construction of a normalized directionally cloned cDNA library from
 adult heart and analysis of 3040 clones by partial sequencing

JOURNAL Genomics 35 (1), 231-235 (1996)

COMMENT 9629752
 Contact: Yusuke Nakamura
 Institute of Medical Science
 University of Tokyo
 4-6-1, Shirokanedai, Minato-ku, Tokyo 108, Japan
 Tel: 81-3-5449-5372
 Fax: 81-3-5449-5433
 Email: yusuke@ims.u-tokyo.ac.jp.

BASE COUNT 24 a 26 c 37 g 17 t

FEATURES

source
 1..104
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="3NHC4544"
 /clone_1lb="Human heart cDNA (Ynakamura)"
 /dev_stage="adult"
 /note="Organ: heart; normalized directionally cloned cDNA
 from adult heart"

Query Match 0.1%; Score 38; DB 27; Length 104;
 Best Local Similarity 100.0%; Pred. No. 0.13;
 Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 28378 GCCTCCGGGTCAAGCGATTCCTCGCTCAGCCTCC 28415

Db 38 GCCTCCGGGTCAAGCGATTCCTCGCTCAGCCTCC 1

RESULT 11
 F24490 80 bp mRNA EST 13-MAY-1999
 LOCUS F24490 HSPD10834 HM3 Homo sapiens cDNA clone s4000013A06, mRNA sequence.

DEFINITION F24490
 VERSION F24490.1 GI:4810116

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Homiidae; Homo.
 1 (bases 1 to 80)

REFERENCE Lanfranchi,G., Muraro,T., Caldera,F., Pacchioni,B., Pallavicini,A.,
 Pandolfo,D., Toppo,S., Trevisan,S., Scarso,S. and Valle,G.
 Identification of 4370 expressed sequence tags from a

3'-end-specific cDNA library of human skeletal muscle by DNA
 sequencing and filter hybridization
 Genome Res. 6 (1), 35-42 (1996)
 96276048

COMMENT On Jun 5, 1998 this sequence version replaced gi:3188834.
 Contact: Valle G.
 CRIBI Biotechnology Centre
 University of Padua
 Via Trieste 75, 35121 Padua, Italy
 ABI Chromatograms and other information are available on WWW at
 http://grup.bio.unipd.it.

FEATURES
 source
 1..80
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="s4000013A06"
 /clone_1lb="HM3"

/sex="female"
 /tissue_type="pectoral muscle (after mastectomy)"
 /note="Vector: pCDNAII (Invitrogen); Site:1: BstXI;
 Site:2: NotI; The library was constructed by G.
 Lanfranchi. This library is not subtracted nor normalized.
 The first strand cDNA was primed with a biotinylated
 oligo-dT-NotI primer
 (5'-biotin-AACCGGCTGAGAGCGCGCTTTTCTTTTCTTTT-3'). The
 ds cDNA was sonicated and size-selected in the range
 350-550 bp. The 3' specific fragments were selected by
 streptavidin coated magnetic beads, ligated to
 non-palindromic BstXI adapters, NotI digested and
 directionally cloned into BstXI-NotI cut pCDNAII vector."

BASE COUNT 18 a 21 c 19 g 22 t

Query Match 0.1%; Score 37; DB 50; Length 80;
 Best Local Similarity 100.0%; Pred. No. 0.3;
 Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 13521 GTGGCTCATGCTGTATCCAGACACTTGGAGGCC 13557

Db 44 GTGGCTCATGCTGTATCCAGACACTTGGAGGCC 80

RESULT 12
 AA078527/c 96 bp mRNA EST 24-SEP-1999
 LOCUS AA078527 7P05B09 Chromosome 7 Placental cDNA library Homo sapiens cDNA clone
 DEFINITION 7P05B09, mRNA sequence.

ACCESSION AA078527
 VERSION AA078527.1 GI:1838015
 KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Homiidae; Homo.
 1 (bases 1 to 96)

REFERENCE Touchman,J.W., Bouffard,G.G., Weintrub,L.A., Idol,J.R., Wang,L.,
 Robbins,C.M., Nussbaum,J.C., Lovett,M. and Green,E.D.
 2006 expressed-sequence tags derived from human chromosome
 7-enriched cDNA libraries
 Genome Res. 7 (3), 281-292 (1997)

JOURNAL 97228905
 MEDLINE
 COMMENT On Sep 12, 1996 this sequence version replaced gi:1288977.
 Contact: Eric D. Green
 Genome Technology Branch
 National Human Genome Research Institute/NIH
 49 Convent Dr., MSC4431, Building 49, Room 2A08, Bethesda, MD 20892
 Tel: 3014020201
 Fax: 3014024735

EMAIL: egreen@hghri.nih.gov
 Plate: 05 row: B column: 09
 Seq primer: -21M3 (ABI).
 Location/Qualifiers

FEATURES


```
source
1. .96
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="7p05B09"
/clone_11b="Chromosome 7 Placental cDNA Library"
/sex="female"
/dev_stage="Placenta obtained at birth (full term)"
/lab_host="E. coli strain DH5 alpha"
/notes="Organ: placenta; Vector: pAMP10; cDNA was generated from cytoplasmic RNA using a mixture of random DNA hexamers and Oligo(dT). From this pool of cDNA, human chromosome 7-enriched cDNA was isolated by direct cDNA selection using chromosome 7 genomic DNA (cosmids). The resulting direct-selected cDNA was cloned into a plasmid vector using a non-directional uracil DNA glycosylase (UDG)-mediated cloning strategy."
```

BASE COUNT 25 a 29 c 20 g 22 t

ORIGIN

Query Match 0.1%; Score 37; DB 28; Length 36;
Best Local Similarity 100.0%; Pred. No. 0.27;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 13518 GTGTGCTCATGCTGATCCACACACTTGGAG 13554
|||||
DB 48 GTGTGCTCATGCTGATCCACACACTTGGAG 12

RESULT 13
AA252633 100 bp mRNA EST 12-MAR-1997
LOCUS AA252633
DEFINITION IMAGE:532504.5' similar to contains Alu repetitive element; contains element MER22 repetitive element; mRNA sequence.

ACCESSION AA252633
VERSION AA252633.1 GI:1887667
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 100)
Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chapell, B., Chissole, S., Dietrich, N., Dubuque, T., Favello, A., Gish, W., Hawkins, M., Hultman, M., Kucaba, T., Lacy, M., Le, M., Le, N., Marais, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L., Rohlfing, T., Schellenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J., Trevasakis, E., Underwood, K., Wohlmann, P., Waterston, R., Wilson, R., and Marra, M.

TITLE Generation and analysis of 280,000 human expressed sequence tags

JOURNAL Genome Res. 6 (9), 807-828 (1996)
MEDLINE 97044478

COMMENT On May 5, 1995 this sequence version replaced g1:798262.
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: estewatson.wustl.edu
This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Putative full length read
The vector to vector length is 101
Seq primer: 28m3 rev1 ET from Amersham.

FEATURES
source
1. .100
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="GDB:5182587"
/db_xref="taxon:9606"
/clone="IMAGE:532504"
/clone_11b="Stratagene hNT neuron (#937233)"
/dev_stage="hNT neurons"

BASE COUNT 21 a 24 c 26 g 28 t 2 others

ORIGIN

Query Match 0.1%; Score 37; DB 30; Length 100;
Best Local Similarity 100.0%; Pred. No. 0.27;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 19129 GAGATCGACACTGCATCCGCGGCGACAGG 19165
|||||
DB 49 GAGATCGACACTGCATCCGCGGCGACAGG 85

RESULT 14
N80946/c 101 bp mRNA EST 02-APR-1996
LOCUS N80946
DEFINITION IMAGE:292298.5' similar to contains Alu repetitive element; mRNA sequence.

ACCESSION N80946
VERSION N80946.1 GI:1243647
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 101)
Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevasakis, E., Waterston, R., Williamson, A., Wohlmann, P., and Wilson, R.

TITLE The WashU-Merck EST Project

JOURNAL Unpublished (1995)
MEDLINE

COMMENT On Apr 14, 1993 this sequence version replaced g1:837971.
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: estewatson.wustl.edu
This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: reverse ET
High quality sequence stop: 67.

FEATURES
source
1. .101
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="GDB:380071"
/db_xref="taxon:9606"
/clone="IMAGE:292298"
/clone_11b="Soares fetal liver spleen INFLS"
/sex="male"
/dev_stage="20 week-post conception fetus"
/lab_host="DH10B (ampicillin resistant)"
/notes="Organ: Liver and Spleen; Vector: p773D (Pharmacia) with a modified polylinker; Site_1: Pac I; Site_2: Eco RI; 1st strand cDNA was primed with a Pac I - oligo(dT) primer [5' AACGTGAGAGATTAATTAAGATTTTCTTTTCTTTT 3'], double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Pac I and cloned into the Pac I and Eco RI sites of the modified p773 vector. Library went through one round of normalization. Library constructed by Bento Soares and M. Fatima Bonaldo."

Job time: 171429 sec

Query Match 0.1%; Score 37; DB 25; Length 101;
 Best Local Similarity 100.0%; Pred. No. 0.27;
 Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24101 CCTGCCACATGGTGAACCCGCTCTCTACTAAAT 24137
 ||||||||||||||||||||||||||||||||||||
 Db 39 CCTGCCACATGGTGAACCCGCTCTCTACTAAAT 3

RESULT 15

AA078709

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

COMMENT

On Sep 12, 1996 this sequence version replaced g1:1293976.

Contact: Eric D. Green

Genome Technology Branch

National Human Genome Research Institute/NIH

49 Convent Dr., MSC4431, Building 49, Room 2A08, Bethesda, MD 20892

Tel: 3014020201

Fax: 3014024735

Email: egreen@hghri.nih.gov

Plate: 01 row: E column: 11

Seq primer: -21M13 (AB1).

Location/Qualifiers

1..74

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="7T01E11"

/clone.lib="Chromosome 7 Thymus cDNA Library"

/sex="mixture of female and male"

/tissue_type="thymus"

/dev_stage="1 week old (male), 11 year old (female)"

/lab_host="E. coli strain DH5 alpha"

/note="Organ: thymus; Vector: PAMP10; cDNA was generated from cytoplasmic RNA using a mixture of random DNA hexamers and oligo(dT). From this pool of cDNA, human chromosome 7-enriched cDNA was isolated by direct cDNA selection using chromosome 7 genomic DNA (cosmids). The resulting direct-selected cDNA was cloned into a plasmid vector using a non-directional uracil DNA glycosylase (UDG)-mediated cloning strategy."

BASE COUNT

19 a

17 c

22 g

16 t

ORIGIN

Query Match

Best Local Similarity

Matches

36; Conservative

0; Mismatches

0; Indels

0; Gaps

0;

QY 13521 GTGGCTCATGCTGTATCCAGCAGCTTTGGAGGC 13556

|||||

Db 39 GTGGCTCATGCTGTATCCAGCAGCTTTGGAGGC 74

|||||

Search completed: March 31, 2000, 02:52:55

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 31, 2000, 06:36:34 ; Search time 13412.9 Seconds

(without alignments)
-6791.568 Million cell updates/sec

Title: US-08-852-495c-1_COPY_165000_195000

Perfect score: 30001
Sequence: 1 AGTGAATATTTATAGAGTCC.....GCACTTTTCATTCTCTCGGT 30001

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 821193 seqs, -1518192014 residues

Word size : 0

Total number of hits satisfying chosen parameters: 345848

Minimum DB seq length: 8

Maximum DB seq length: 105

Post-processing: Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba1:*
2: gb_ba2:*
3: gb_om:*
4: gb_om:*
5: gb_pat:*
6: gb_ph:*
7: gb_pl1:*
8: gb_pl2:*
9: gb_pl1:*
10: gb_pl2:*
11: gb_pl3:*
12: gb_ro:*
13: gb_sts:*
14: gb_sy:*
15: gb_un:*
16: gb_vl:*
17: em_fun:*
18: em_hum1:*
19: em_hum2:*
20: em_in:*
21: em_om:*
22: em_or:*
23: em_ov:*
24: em_pat:*
25: em_ph:*
26: em_pl:*
27: em_ro:*
28: em_sts:*
29: em_sy:*
30: em_un:*
31: em_vl:*
32: gb_hcgl:*
33: gb_hcgl:*
34: gb_hcgl:*
35: gb_hcgl:*
36: em_ba1:*
37: em_ba2:*
38: em_hum3:*
39: em_hum4:*
40: gb_pt4:*
41: gb_hcgl3:*
42: gb_hcgl4:*
43: gb_hcgl5:*

44: gb_hcgl6:*
45: gb_hcgl7:*
46: em_hcgl1:*
47: em_hcgl2:*
48: em_hcgl3:*
49: em_hum5:*
50: gb_pl3:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	54	0.2	60	5	I31332 Sequence 24
2	50	0.2	91	5	I31254 Sequence 16
3	50	0.2	92	5	I31518 Sequence 43
4	46	0.2	46	5	I31437 Sequence 34
5	42	0.1	56	5	I31153 Sequence 65
6	41	0.1	50	5	I31467 Sequence 37
7	40	0.1	40	5	A68622 Sequence 2
8	36	0.1	40	5	A68621 Sequence 1
9	36	0.1	40	5	A68621 Sequence 1
10	36	0.1	103	9	HUMALCE221
11	35	0.1	35	5	A25212
12	35	0.1	35	5	A25212 Inter-Alu s
13	35	0.1	35	24	E09140
14	35	0.1	35	24	E09140 Synthetic D
15	35	0.1	91	13	L30244 Human STS U
16	35	0.1	102	13	G37854 PF2688 Plas
17	34	0.1	63	13	L30829 Human STS U
18	34	0.1	72	13	G31952 MCM92 Chick
19	34	0.1	75	3	BTGA45
20	34	0.1	80	9	HUMBRKEAB
21	34	0.1	80	9	HUMBRKEAB
22	34	0.1	88	3	ORU15702
23	34	0.1	95	13	A0026531
24	34	0.1	95	13	A0026788
25	33	0.1	33	5	I31400
26	33	0.1	40	5	I31263 Sequence 17
27	33	0.1	46	5	I31159
28	33	0.1	51	5	I31405
29	33	0.1	54	5	AR002288
30	33	0.1	54	5	AR053139
31	33	0.1	60	34	DRORS1
32	33	0.1	62	5	I07633
33	33	0.1	79	10	S73203
34	33	0.1	82	9	HUMRSCB
35	33	0.1	87	9	HUWL
36	33	0.1	87	12	MMVIMV24
37	33	0.1	87	13	G31698
38	33	0.1	90	12	MMDNDS22
39	33	0.1	91	12	MMDNDS1
40	33	0.1	95	12	MMDNDS2
41	33	0.1	95	34	AG2H79
42	33	0.1	97	3	SHPMCR
43	33	0.1	97	9	HUWLDLRA2
44	33	0.1	97	9	HUWLDLRLD
45	33	0.1	98	34	AG2H815

ALIGNMENTS

RESULT 1
LOCUS I31332 60 bp DNA
DEFINITION Sequence 244 from patent US 5582979.
ACCESSION I31332
VERSION I31332.1 GI:1822123 PAT 06-FEB-1997

KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	FEATURES	BASE COUNT	ORIGIN
Unknown.	Unknown.	Unknown.	Unclassified.	1 (bases 1 to 60)	Weber,J.T.	Length polymorphisms in (dc-da).sub.n..(dg-dt).sub.n sequences and method of using the same	Patent: US 5582979-A 244 10-DEC-1996;	18 a	12 t
					Location/Qualifiers		1..60	30 c	0 g
							"/organism="unknown"		

[illegible]

RESULT	2				
LOCUS	131254				
DEFINITION	I31254	91 bp	DNA	PAT	06-FEB-1997
ACCESSION	I31254	Sequence 166 from patent US 5582979.			
VERSION	I31254.1	G1:1822045			
KEYWORDS	.				
SOURCE	Unknown.				
ORGANISM	Unknown.				
REFERENCE	Unclassified.				
AUTHORS	1 (bases 1 to 91)				
TITLE	Weber,J.L.				
JOURNAL	Length polymorphisms in (dc-da).sub.n.(dg-dt).sub.n sequences and method of using the same				
FEATURES	Patent: US 5582979-A 166 10-DEC-1996;				
Source	Location/Qualifiers				
	1..91				
	/organism="unknown"				
BASE COUNT	20 a	43 c	1 g	27 t	
ORIGIN					

	Query Match	Similarity	0.2%	Score	50:	DB	5:	Length	91:
	Best Local	Similarity	100.0%	Pred.	No.	3-8e-13:			
Matches	50:	Conservative	0:	Mismatches	0:	Indels	0:	Gaps	0:
QY	28932	CTCTCTCTCTCTCTCTCTCTCTCACACACACACACACACACA	28981						
DG	30	CTCTCTCTCTCTCTCTCTCTCTCACACACACACACACACACA	79						

RESULT	3				
LOCUS	131518				
DEFINITION	Sequence 430 from patent US 5582979.			PAT	06-FEB-1997
ACCESSION	131518				
VERSION	131518.1		GI:1822309		
KEYWORDS					
SOURCE	Unknown.				
ORGANISM	Unknown.				
REFERENCE	Unclassified.				
AUTHORS	1 (bases 1 to 92)				
TITLE	Weber, J.L.				
JOURNAL	length polymorphisms in (dc-da).sub.n.(dc-dt).sub.n sequences and				
FEATURES	method of using the same				
source	Patent: US 5582979-A 430 10-DEC-1996;				
	Location/Qualifiers				
	1..92				

BASE COUNT	21 a	43 c	1 g	27 t
ORIGIN	/organism="unknown"			

Query Match	Similarity	100.0%	Score	50	DB	5	Length	92	
Best Local			Pred.	No.	3	8e-13			
Matches	50	Conservative	0	Mismatches	0	Indels	0	Gaps	0
QY	28932	CTCTCTCTCTCTCTCTCTCTCTCAACACACACACACACACACACACACA	28981						
DB	30	CTCTCTCTCTCTCTCTCTCTCTCAACACACACACACACACACACACACA	79						

RESULT		4		PAT	06-FEB-1997
I11437					
LOCUS	I11437	46 bp	DNA		
DEFINITION	Sequence 349 from patent US 5582979.				
ACCESSION	I11437				
VERSION	I11437.1	GI:1822228			
KEYWORDS	.				
SOURCE	Unknown.				
ORGANISM	Unknown.				
REFERENCE	Unclassified.				
AUTHORS	1 (bases 1 to 46)				
TITLE	Weber,J.L.				
JOURNAL	length polymorphisms in (dc-da).sub.n.(dc-dr).sub.n sequences and method of using the same				
FEATURES	Patent: US 5582979-A 349 10-DEC-1996;				
Source	Location/Qualifiers				
	1..46				
	/organism="unknown"				
BASE COUNT	12 a	0 c	0 g	34 t	
ORIGIN					

```

Query Match      0.2%  Score 46:  DB 5:  Length 46:
Best Local Similarity 100.0%:  Pred. NO. 4e-11:
Matches 46:  Conservative 0:  Mismatches 0:  Indels 0:  Gaps 0:

QY 21515 ATTATATTATTATTATTATTATTATTATTATTATTATTAT 21560
Db 1 ATTATATTATTATTATTATTATTATTATTATTATTATTATTT 46

```

	RESULT	5					
	LOCUS	131153					
	DEFINITION	Sequence 65 from patent US 5582979.	PAT			06-FEB-1997	
	ACCESSION	U31155					
	VERSION	U31155.1	GI:1821944				
	KEYWORDS						
	SOURCE	Unknown.					
	ORGANISM	Unkown.					
	REFERENCE	Unclassified.					
	AUTHORS	1 (bases 1 to 56)					
	TITLE	Weber J.L.					
JOURNAL.		length polymorphisms in (dc-da).sub.n.(dc-dr).sub.n sequences and method of using the same					
FEATURES		Patent: US 5582979-A 65 10-DEC-1996;					
	Source	Location/Qualifiers					
		1..56					
		/organism="unknown"					
BASE COUNT	23	a	28	c	0	g	5 t
ORIGIN							

Query Match	0.18;	Score 42;	DB 5;	Length 56;
Best Local Similarity	100.0%;	Pred. No. 3,7e-09;		
Matches 42;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
OY 28944 CTCTCTCTCTACACACACACACACACACACACACA				

Db 1 CTCCTCTCTCACACACACACACACACACACACACA 42

RESULT 6
LOCUS I31467 50 bp DNA PAT 06-FEB-1997
DEFINITION Sequence 379 from patent US 5582979.
ACCESSION I31467
VERSION I31467.1 GI:1822258
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 50)
AUTHORS Weber,J.L.
TITLE Length polymorphisms in (dc-da).sub.n.(dg-dt).sub.n sequences and method of using the same
JOURNAL Patent: US 5582979-A 379 10-DEC-1996;
FEATURES Location/Qualifiers
source 1..50 /organism="unknown"
BASE COUNT 20 a 25 c 0 g 5 t
ORIGIN

Query Match 0.1%; Score 41; DB 5; Length 50;
Best Local Similarity 100.0%; Pred. No. 1.2e-08;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 28945 TCCTCTCTCTCACACACACACACACACACACACA 28985
Db 1 TCCTCTCTCTCACACACACACACACACACACACACA 41

RESULT 7
LOCUS A68622 40 bp DNA PAT 06-MAY-1999
DEFINITION Sequence 2 from Patent WO9801573.
ACCESSION A68622
VERSION A68622.1 GI:4759649
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 40)
AUTHORS Resnick,M.A., Larionov,V.L., Koupriina,N.Y. and Perkins,E.L.
TITLE TRANSFORMATION-ASSOCIATED RECOMBINATION CLONING
JOURNAL Patent: WO 9801573-A 15-JAN-1998;
FEATURES Location/Qualifiers
source 1..40 /organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 9 a 8 c 19 g 4 t
ORIGIN

Query Match 0.1%; Score 40; DB 5; Length 40;
Best Local Similarity 100.0%; Pred. No. 3.8e-08;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3376 CCGCCTCCCGGCTCAAGCGATTCTCCGCTCAGCCTCC 3415
Db 40 CCGCCTCCCGGCTCAAGCGATTCTCCGCTCAGCCTCC 1

RESULT 8
LOCUS A68621 40 bp DNA PAT 06-MAY-1999
DEFINITION Sequence 1 from Patent WO9801573.
ACCESSION A68621
VERSION A68621.1 GI:4759648
KEYWORDS

SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 40)
AUTHORS Resnick,M.A., Larionov,V.L., Koupriina,N.Y. and Perkins,E.L.
TITLE TRANSFORMATION-ASSOCIATED RECOMBINATION CLONING
JOURNAL Patent: WO 9801573-A 15-JAN-1998;
FEATURES Location/Qualifiers
source 1..40 /organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 7 a 12 c 13 g 8 t
ORIGIN

Query Match 0.1%; Score 36; DB 5; Length 40;
Best Local Similarity 100.0%; Pred. No. 3.6e-06;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19729 TGGCTCAGCCTGTAATCCAGCACCTTTGGAGGCC 19764
Db 3 TGGCTCAGCCTGTAATCCAGCACCTTTGGAGGCC 38

RESULT 9
LOCUS A68621 40 bp DNA PAT 06-MAY-1999
DEFINITION Sequence 1 from Patent WO9801573.
ACCESSION A68621
VERSION A68621.1 GI:4759648
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 40)
AUTHORS Resnick,M.A., Larionov,V.L., Koupriina,N.Y. and Perkins,E.L.
TITLE TRANSFORMATION-ASSOCIATED RECOMBINATION CLONING
JOURNAL Patent: WO 9801573-A 15-JAN-1998;
FEATURES Location/Qualifiers
source 1..40 /organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 7 a 12 c 13 g 8 t
ORIGIN

Query Match 0.1%; Score 36; DB 5; Length 40;
Best Local Similarity 100.0%; Pred. No. 3.6e-06;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21783 GCCTCCCAAGTCTGGGATTACAGGCGTGAGCCAC 21818
Db 37 GCCTCCCAAGTCTGGGATTACAGGCGTGAGCCAC 2

RESULT 10
LOCUS HUMALCE221 103 bp ss-RNA PRI 15-APR-1994
DEFINITION Human carcinoma cell-derived Alu RNA transcript, clone CE221.
ACCESSION M87896
VERSION M87896.1 GI:174874
KEYWORDS Alu repeat.
SOURCE Homo sapiens male embryo carcinoma cDNA to other RNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 103)
AUTHORS Simelett,D., Richter,C., Deragon,J.-M. and Labuda,D.
TITLE Alu RNA transcripts in human embryonal carcinoma cells. Model of post-transcriptional selection of master sequences
JOURNAL J. Mol. Biol. (1992) In press

FEATURES
source Location/Qualifiers
1. 103
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_line="Ntera2D1"
/dev_stage="embryo"
/sex="male"
/tissue_type="carcinoma"
BASE COUNT 25 a 27 c 33 g 18 t
ORIGIN

Query Match
Best Local Similarity 0.1%; Score 36; DB 9; Length 103;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3372 ACCTCCGCTCCCGGGTCAAGCGATTCTGCTGCT 3407
DB 69 ACCTCCGCTCCCGGGTCAAGCGATTCTGCTGCT 34

RESULT 11
A25212 35 bp DNA PAT 11-APR-1995
LOCUS A25212 Inter-Alu specific primer DNA (pdj33) from patent WO9213101.
DEFINITION A25212
ACCESSION A25212
VERSION A25212.1 GI:904592
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 35)
AUTHORS
TITLE METHOD OF DETECTING DNA SEQUENCE VARIATION
JOURNAL Patent: WO 9213101-A 3 06-AUG-1992;
FEATURES Location/Qualifiers
source
1. 35
/organism="synthetic construct"
/db_xref="taxon:32630"
BASE COUNT 8 a 10 c 11 g 6 t
ORIGIN

Query Match
Best Local Similarity 0.1%; Score 35; DB 5; Length 35;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 21783 GCCTCCCAAAGTGTGGATTACAGCGGTGAGCCA 21817
DB 1 GCCTCCCAAAGTGTGGATTACAGCGGTGAGCCA 35

RESULT 12
A25212/c 35 bp DNA PAT 11-APR-1995
LOCUS A25212 Inter-Alu specific primer DNA (pdj33) from patent WO9213101.
DEFINITION A25212
ACCESSION A25212
VERSION A25212.1 GI:904592
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 35)
AUTHORS
TITLE METHOD OF DETECTING DNA SEQUENCE VARIATION
JOURNAL Patent: WO 9213101-A 3 06-AUG-1992;
FEATURES Location/Qualifiers
source
1. 35
/organism="synthetic construct"
/db_xref="taxon:32630"
BASE COUNT 8 a 10 c 11 g 6 t
ORIGIN

Query Match
Best Local Similarity 0.1%; Score 35; DB 5; Length 35;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 19729 TGGCTCAGCGCTGTATCCACGACTTTGGAGGC 19763
DB 35 TGGCTCAGCGCTGTATCCACGACTTTGGAGGC 1

RESULT 13
E09140 standard; DNA; UNC; 35 BP.
ID E09140
XX
AC E09140;
XX
SV E09140.1
XX
XX 07-OCT-1997 (Rel. 52, Created)
DT 08-OCT-1997 (Rel. 52, Last updated, Version 1)
XX
DE Synthetic DNA for Alu specific primer.
XX
XX JP 1995115999-A/3.
XX
OS unidentified.
OC unclassified.
XX
XX [1]
RN 1-35
RP Andoreasu H.A., Yan F.;
RA "DETECTING METHOD FOR DNA ARRANGEMENT VARIATION";
RT Patent number JP 1995115999-A/3, 09-MAY-1995.
RL INGENIT BV.
XX
XX OS None
CC Artificial sequences.
CC OC JP 1995115999-A/3
CC PN 09-MAY-1995
CC PD 22-MAY-1992 JP 1992130668
CC PE PI ANDOREASU HERARUDOSU AITSUTERURINDEN, YAN FUEIKU
CC PC C1201/68,C12N15/00,G01N27/447,G01N27/447;
CC CC strandedness: Single;
CC CC topology: Linear;
CC CC Key Location/Qualifiers
CC FH source 1. 35
CC FT /organism="Artificial sequences"
CC FT misc_feature 1. 35
CC FT /note="Alu specific primer"
XX
FH Key Location/Qualifiers
FH FT source 1. 35
FH FT /db_xref="taxon:32644"
FH FT /organism="unidentified"
XX
SQ Sequence 35 BP; 8 A; 10 C; 11 G; 6 T; 0 other;

Query Match
Best Local Similarity 0.1%; Score 35; DB 24; Length 35;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 21783 GCCTCCCAAAGTGTGGATTACAGCGGTGAGCCA 21817
DB 1 GCCTCCCAAAGTGTGGATTACAGCGGTGAGCCA 35

RESULT 14
E09140/c standard; DNA; UNC; 35 BP.
ID E09140
XX

AC E09140;
XX
SV E09140.1
DT 07-OCT-1997 (Rel. 52, Created)
DT 08-OCT-1997 (Rel. 52, Last updated, Version 1)
XX
DE Synthetic DNA for Alu specific primer.
XX JP 1995115999-A/3.
XX
XX
OS unidentified
OC unclassified.
XX
XX
RN 1-35
RP 1-35
RA Andorea H.A., Van F.;
RT "DETECTING METHOD FOR DNA ARRANGEMENT VARIATION";
RL Patent number JP 1995115999-A/3, 09-MAY-1995.
RL INGENIT BV.
XX
XX
OS None
CC Artificial sequences.
CC PN JP 1995115999-A/3
CC PD 09-MAY-1995
CC PF 22-MAY-1992 JP 1992130668
CC PI ANOREASU HERARDODUSU AITSUTERURINDEN, YAN FUEIKU
CC PC C12Q1/68, C12N15/00, G01N27/447, G01N27/447;
CC CC strandedness: Single;
CC CC topology: Linear;
CC CC Key Location/Qualifiers
CC CC FT 1..35 /organism="Artificial sequences"
CC CC FT misc.feature 1..35
CC CC FT /note="Alu specific primer"
FH Key Location/Qualifiers
FH FT source 1..35 /db_xref="taxon:32644"
FH FT /organism="unidentified"
XX
XX
SO Sequence 35 Bp: 8 A; 10 C; 11 G; 6 T; 0 other;

Query Match 0.1%; Score 35; DB 24; Length 35;
Best Local Similarity 100.0%; Pred. No. 1.1e-05;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19729 TGGCTCAGCGCTGTATCCAGACATTGGAGGC 19763
DB 35 TGGCTCAGCGCTGTATCCAGACATTGGAGGC 1

RESULT 15
LOCUS HUMUT8164A 91 bp DNA STS 29-DEC-1994
DEFINITION Human STS UT8164, 5' primer bind, sequence tagged site.
ACCESSION L30244
VERSION L30244.1 GI:605447
KEYWORDS STS; PCR primer; STS sequence; microsatellite DNA; microsatellite marker; sequence tagged site; tetranucleotide repeat.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 91)
AUTHORS Gerken,S.C., Matsunami,N., Plaetke,R., Albertsen,E., Ballard,L., Wells,R., Lawrence,E., Moore,M., Holik,P.R., Carlson,M., Zhao,X., Robertson,M., Bradley,P., Elsner,T., Tinney,A., Lalouel,J.-M. and White,R.
TITLE Genetic and physical mapping of simple sequence repeat containing

JOURNAL
COMMENT
sequence tagged sites from the human genome
Unpublished (1994)
Submitted by: Utah Center for Human Genome Research University of
Utah, Dept. of Human Genetics
2160 Eccles Institute of Human Genetics
Salt Lake City, UT 84112
e-mail: stgcorona.med.utah.edu
Primer A: AGAGCTGCAGTGAACCA
Primer B: TTTTCCCCCTCTACTACT
End to Label: Primer B

PCR Profile:
Initial Denaturation: 94C 300sec
Cycles Denaturation Annealing Extension 5 94
C 10 sec. 56 C 10 sec. 72 C 20 sec. 30
52 C 10 sec. 72 C 20 sec. Mg++: 1.50 mM
Gel: Acrylamide 7%, Formamide 32%, Urea 348
Alleles: 1.

FEATURES
source Location/Qualifiers
1..91
/organism="Homo sapiens"
/db_xref="taxon:9606"
10..28
primer_bind /evidence=experimental 13 t
BASE COUNT 35 a 20 c 23 g
ORIGIN

Query Match 0.1%; Score 35; DB 13; Length 91;
Best Local Similarity 100.0%; Pred. No. 9.9e-06;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3317 GAGCTCAGCTGTGTCACCCAGCTGGAGTGCAGTG 3351
DB 71 GAGCTCAGCTGTGTCACCCAGCTGGAGTGCAGTG 37

Search completed: March 31, 2000, 17:25:40
Job time: 223289 sec

Matches 63; Conservatively

Query Match	0.24;	Score 63;	DB 1;	Length 88;
Best Local Similarity	100.04;	Pred. No. 5,8e-08;		
Matches 63;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

21514 TATTATTATTATTATTATTATTATTATTATTATTATTATTATTATTGACACGAG 21573

DE Alu PCR primer 1.
KM PCR; primer; amplification; Alu repeat sequence; vector;
KW circular yeast artificial chromosome; YAC; ss.
OS Synthetic.
OS Saccharomyces sp.
PN WO9801573-A1.
PD 15-JAN-1998.
PF 09-JUL-1996; U11478.
PR 09-JUL-1996; WO-U11478.
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
PI Kourilna NY, Lariouov VL, Perkins EL, Resnick MA;
DR WPI; 98-110234/10.
PT Preparation of yeast artificial chromosomes - by in vivo
PT recombination using vector comprising yeast centromere, marker,
PT yeast telomere and nucleic acid for recombination
PS Example 1; Page 45; 117pp; English.
CC This is the nucleotide sequence for the PCR primer used in the
CC amplification of the Alu repeat sequence, which is used to
CC demonstrate the processes described in the invention. It involves
CC the creation and use of circular yeast artificial chromosome (YAC)
CC to selectively clone specific nucleic acids from a background of
CC mixed nucleic acids by introducing the vector(s) into E. coli cells.
CC They can be used to rapidly isolate human DNA where only a part of the
CC sequence of DNA is known. Using the methods large fragments of DNA can
CC be easily cloned and analysed.
SQ Sequence 40 BP; 7 A; 12 C; 13 G; 8 T;

Query Match 0.1%; Score 36; DB 1; Length 40;
Best Local Similarity 100.0%; Pred. No. 0.56;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 19729 TGCGTCACGCGCTGTATCCAGCATTGGAGGCC 19764
|||||
DB 3 TGCGTCACGCGCTGTATCCAGCATTGGAGGCC 38

RESULT 15
VI9044/C
ID VI9044 standard; DNA; 40 BP.
AC VI9044;
DT 28-JUL-1998 (first entry)
DE Alu PCR primer 1.
KM PCR; primer; amplification; Alu repeat sequence; vector;
KW circular yeast artificial chromosome; YAC; ss.
OS Synthetic.
OS Saccharomyces sp.
PN WO9801573-A1.
PD 15-JAN-1998.
PF 09-JUL-1996; U11478.
PR 09-JUL-1996; WO-U11478.
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
PI Kourilna NY, Lariouov VL, Perkins EL, Resnick MA;
DR WPI; 98-110234/10.
PT Preparation of yeast artificial chromosomes - by in vivo
PT recombination using vector comprising yeast centromere, marker,
PT yeast telomere and nucleic acid for recombination
PS Example 1; Page 45; 117pp; English.
CC This is the nucleotide sequence for the PCR primer used in the
CC amplification of the Alu repeat sequence, which is used to
CC demonstrate the processes described in the invention. It involves
CC the creation and use of circular yeast artificial chromosome (YAC)
CC to selectively clone specific nucleic acids from a background of
CC mixed nucleic acids by introducing the vector(s) into E. coli cells.
CC They can be used to rapidly isolate human DNA where only a part of the
CC sequence of DNA is known. Using the methods large fragments of DNA can
CC be easily cloned and analysed.
SQ Sequence 40 BP; 7 A; 12 C; 13 G; 8 T;

Query Match 0.1%; Score 36; DB 1; Length 40;
Best Local Similarity 100.0%; Pred. No. 0.56;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 21783 GCCTCCCAAGTGTGATACAGCGGTAGCCAC 21818
|||||
DB 37 GCCTCCCAAGTGTGATACAGCGGTAGCCAC 2

Search completed: March 31, 2000, 19:25:47
Job time: 228945 sec

Mon Apr 3 08:24:10 2000

us-08-852-495c-1_copy_165000_195000.png

Page 7

Mon Apr 3 08:24:11 2000

us-08-852-495c-1_copy_165000_195000.rml

Page 1

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 31, 2000, 10:04:17 ; Search time 340.38 Seconds
(Without alignments)
10552.365 Million cell updates/sec

Title: US-08-852-495c-1_COPY_165000_195000

Perfect score: 30001
Sequence: 1 AGTGAAATTTATAGAGTCC.....GCAGTTTCATTGTGTGTGT 30001.

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 214294 seqs, 59861574 residues

Word size : 0

Total number of hits satisfying chosen parameters: 349160

Minimum DB seq length: 8
Maximum DB seq length: 105

Post-processing: Listing first 45 summaries

Database : Issued_Patents_NA.*
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq.*
3: /cgn2_6/ptodata/1/ina/5C.COMB.seq.*
4: /cgn2_6/ptodata/1/ina/5D.COMB.seq.*
5: /cgn2_6/ptodata/1/ina/6.COMB.seq.*
6: /cgn2_6/ptodata/1/ina/PTUS9.COMB.seq.*
7: /cgn2_6/ptodata/1/ina/backfile1.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	54	0.2	60	1	US-08-222-177A-244 Sequence 244, App
2	50	0.2	91	1	US-08-222-177A-166 Sequence 166, App
3	50	0.2	92	1	US-08-222-177A-430 Sequence 430, App
4	46	0.2	46	1	US-08-222-177A-349 Sequence 349, App
5	45	0.1	56	1	US-08-222-177A-65 Sequence 65, App1
6	41	0.1	50	1	US-08-222-177A-379 Sequence 379, App
7	33	0.1	33	1	US-08-222-177A-312 Sequence 312, App
8	33	0.1	40	1	US-08-222-177A-175 Sequence 175, App
9	33	0.1	46	1	US-08-222-177A-71 Sequence 71, App1
10	33	0.1	51	1	US-08-222-177A-317 Sequence 317, App
11	33	0.1	54	2	US-08-469-802B-27 Sequence 27, App1
12	33	0.1	54	3	US-08-267-803B-45 Sequence 45, App1
13	33	0.1	62	7	5506118-3 Patent No. 5506118
14	32	0.1	33	1	US-08-222-177A-59 Sequence 59, App1
15	32	0.1	33	1	US-08-222-177A-218 Sequence 218, App
16	32	0.1	34	1	US-08-222-177A-110 Sequence 110, App
17	32	0.1	34	1	US-08-222-177A-112 Sequence 112, App
18	32	0.1	34	1	US-08-222-177A-180 Sequence 180, App
19	32	0.1	34	1	US-08-222-177A-192 Sequence 192, App
20	32	0.1	34	1	US-08-222-177A-322 Sequence 322, App
21	32	0.1	35	1	US-08-222-177A-77 Sequence 77, App1
22	32	0.1	35	1	US-08-222-177A-189 Sequence 189, App
23	32	0.1	36	1	US-08-222-177A-104 Sequence 104, App
24	32	0.1	36	1	US-08-222-177A-358 Sequence 358, App
25	32	0.1	37	1	US-08-222-177A-140 Sequence 140, App
26	32	0.1	38	1	US-08-222-177A-198 Sequence 198, App

27	32	0.1	38	1	US-08-222-177A-397 Sequence 397, App
28	32	0.1	39	1	US-08-222-177A-137 Sequence 137, App
29	32	0.1	40	1	US-08-222-177A-119 Sequence 119, App
30	32	0.1	40	1	US-08-222-177A-400 Sequence 400, App
31	32	0.1	40	1	US-08-222-177A-403 Sequence 403, App
32	32	0.1	41	1	US-08-222-177A-74 Sequence 74, App1
33	32	0.1	41	1	US-08-222-177A-183 Sequence 183, App
34	32	0.1	41	1	US-08-222-177A-232 Sequence 232, App
35	32	0.1	42	1	US-08-222-177A-53 Sequence 53, App1
36	32	0.1	42	1	US-08-222-177A-229 Sequence 229, App
37	32	0.1	43	1	US-08-222-177A-370 Sequence 370, App
38	32	0.1	43	1	US-08-222-177A-388 Sequence 388, App
39	32	0.1	44	1	US-08-222-177A-195 Sequence 195, App
40	32	0.1	44	1	US-08-222-177A-241 Sequence 241, App
41	32	0.1	44	1	US-08-222-177A-249 Sequence 249, App
42	32	0.1	45	1	US-08-222-177A-157 Sequence 157, App
43	32	0.1	45	1	US-08-222-177A-203 Sequence 203, App
44	32	0.1	45	1	US-08-222-177A-364 Sequence 364, App
45	32	0.1	46	1	US-08-222-177A-101 Sequence 101, App

ALIGNMENTS

RESULT 1
US-08-222-177A-244
Sequence 244, Application US/08222177A
Patent No. 5582979
GENERAL INFORMATION:
APPLICANT: Weber, James L.
TITLE OF INVENTION: LENGTH POLYMORPHISMS IN
TITLE OF INVENTION: (dc-da)n.(dg-dt)n SEQUENCES AND METHODS OF USING SAME
NUMBER OF SEQUENCES: 460
CORRESPONDENCE ADDRESS:
ADDRESSEE: Demilt Ross & Stevens, S.C.
STREET: 8000 Excelsior Drive, Suite 401
CITY: Madison
STATE: Wisconsin
COUNTRY: USA
ZIP: 53717-1914
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/222,177A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/341,562
FILING DATE: 21-APR-1989
ATTORNEY/AGENT INFORMATION:
NAME: Sara, Charles S.
REGISTRATION NUMBER: 30,492
REFERENCE/DOCKET NUMBER: 09865.601
TELECOMMUNICATION INFORMATION:
TELEPHONE: (608) 831-2100
TELEFAX: (608) 831-2106
TELEX:
INFORMATION FOR SEQ ID NO: 244:
SEQUENCE CHARACTERISTICS:
LENGTH: 60 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
IMMEDIATE SOURCE:
CLONE: mfd67rs
US-08-222-177A-244
Query Match 0.2%; Score 54; DB 1; Length 60;


```

1  APPLICATION NUMBER:  US 07/341,562
2  FILING DATE:  21-APR-1989
3  ATTORNEY/AGENT INFORMATION:
4  NAME:  Sara, Charles S
5  REGISTRATION NUMBER:  30,492
6  REFERENCE/POCKET NUMBER:  03865.601
7  TELECOMMUNICATION INFORMATION:
8  TELEPHONE:  (608) 831-2100
9  TELEFAX:  (608) 831-2106
10 TELEX:
11 INFORMATION FOR SEQ ID NO:  349:
12 SEQUENCE CHARACTERISTICS:
13 LENGTH:  46 base pairs
14 TYPE:  nucleic acid
15 STRANDEDNESS:  double
16 TOPOLOGY:  linear
17 MOLECULE TYPE:  DNA (genomic)
18 IMMEDIATE SOURCE:
19 CLONE:  mid101rs
20
21 US-08-222-177A-349

```

Query Match	0.28;	Score 46;	DB 1;	Length 46;
Best Local Similarity	100.0%;	Pred. No. 3.9e-06;		
Matches 46;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

[illegible]

```

1 RESULT 5
2 US-08-222-177A-65
3 : Sequence 65, Application US/08222177A
4 : Patent No. 5582979
5 :
6 : GENERAL INFORMATION:
7 :
8 : APPLICANT: Weber, James L.
9 :
10 : TITLE OF INVENTION: LENGTH POLYMORPHISMS IN
11 : TITLE OF INVENTION: (dc-da)n.(dc-dt)n SEQUENCES AND METHODS OF USING SAME
12 : NUMBER OF SEQUENCES: 460
13 :
14 : CORRESPONDENCE ADDRESS:
15 : ADDRESSEE: Dewitt Ross & Stevens, S.C.
16 : STREET: 8000 Excelsior Drive, Suite 401
17 : CITY: Madison
18 : STATE: Wisconsin
19 : COUNTRY: USA
20 : ZIP: 53717-1914
21 :
22 : COMPUTER READABLE FORM:
23 : MEDIUM TYPE: Floppy disk
24 :
25 : COMPUTER: IBM PC compatible
26 : OPERATING SYSTEM: PC-DOS/MS-DOS
27 : SOFTWARE: Patentln Release #1.0, Version #1.25
28 :
29 : CURRENT APPLICATION DATA:
30 : APPLICATION NUMBER: US/08/222.177A
31 : FILING DATE:
32 :
33 : CLASSIFICATION: 435
34 :
35 : PRIOR APPLICATION DATA:
36 : APPLICATION NUMBER: US 07/341.562
37 : FILING DATE: 21-APR-1989
38 :
39 : ATTORNEY/AGENT INFORMATION:
40 : NAME: Sara, Charles S.
41 : REGISTRATION NUMBER: 30,492
42 : REFERENCE/DOCKET NUMBER: 09865.601
43 : TELECOMMUNICATION INFORMATION:
44 : TELEPHONE: (608) 831-2100
45 : TELEFAX: (608) 831-2106
46 :
47 : TELEX:
48 :
49 : INFORMATION FOR SEQ ID NO: 65:
50 :
51 : SEQUENCE CHARACTERISTICS:
52 : LENGTH: 56 base pairs
53 : TYPE: nucleic acid
54 : STRANDEDNESS: double
55 : TOPOLOGY: linear
56 :
57 :
58 :
59 :
60 :

```

```

; MOLECULE TYPE: DNA (genomic)
; IMMEDIATE SOURCE:
; CLONE: mfd5rs
;
US-08-222-177A-65

```

```
Query Match      0.1%; Score 42; DB 1; Length 56;
Best Local Similarity 100.0%; Pred. No. 7e-05;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

Qy 28944 CTCTCTCTCTACACACACACACACACACACA 28985
      |||||
Db 1 CTCTCTCTCTACACACACACACACACACACA 42

```

RESULT 6
US-08-222-177A-379
Sequence 379, Application US/08222177A

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

COUNTRY: USA
ZIP: 53717-1914
COMPUTER READABLE F

```

1  COMPUTER READABLE FORM:
2  MEDIUM TYPE: Floppy disk
3  COMPUTER: IBM PC compatible
4  OPERATING SYSTEM: PC-DOS/MS-DOS
5  SOFTWARE: PatentIn Release #1.0, Version #1.25
6  CURRENT APPLICATION DATA: 00.000.000

```

APPLICATION NUMBER: US/08/222,177A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/341,562
PUBLICATION NO: 2000

APPLICATION NUMBER: US 07/341,562
FILING DATE: 21-APR-1989
ATTORNEY/AGENT INFORMATION:
NAME: Sara, Charles S.
REGISTRATION NUMBER: 30,492
REFERENCE/DOCKET NUMBER: 09865.600

```

? REFERENCE/DOCKET NUMBER: 09865.601
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (608) 831-2100
? TELEFAX: (608) 831-2106
? TELE:
? INFORMATION FOR SEQ ID NO: 379:

```

```

; INFORMATION FOR SEQ ID NO: 379
;
; SEQUENCE CHARACTERISTICS:
;     LENGTH: 50 base pairs
;     TYPE: nucleic acid
;     STRANDEDNESS: double
;     TOPOLOGY: linear

```

```

; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; IMMEDIATE SOURCE:
; CLONE: mfd120rs
;
US-08-222-177A-379

```

Query Match	0.1%;	Score 41;	DB 1;	Length 50;
Best Local Similarity	100.0%;	Pred. No. 0.00015;		
Matches 41;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

```

Qy 28945 TCTCTCTCTCACAACACACACACACACACACACACA 28985
      |||||
Db 1 TCTCTCTCTCACAACACACACACACACACACACACA 41

```

RESULT 7

US-08-222-177A-312
Sequence 312, Application US/08222177A
Patent No. 5582979
GENERAL INFORMATION:
APPLICANT: Weber, James L.
TITLE OF INVENTION: LENGTH POLYMORPHISMS IN
SEQUENCES AND METHODS OF USING SAME
NUMBER OF SEQUENCES: 460
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dewitt Ross & Stevens, S.C.
STREET: 8000 Excelsior Drive, Suite 401
CITY: Madison
STATE: Wisconsin
COUNTRY: USA
ZIP: 53717-1914
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/222,177A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/341,562
FILING DATE: 21-APR-1989
ATTORNEY/AGENT INFORMATION:
NAME: Sara, Charles S.
REGISTRATION NUMBER: 30,492
REFERENCE/DOCKET NUMBER: 09865,601
TELECOMMUNICATION INFORMATION:
TELEPHONE: (608) 831-2100
TELEFAX: (608) 831-2106
TELEX:
INFORMATION FOR SEQ ID NO: 312:
SEQUENCE CHARACTERISTICS:
LENGTH: 33 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
MOLECULE TYPE: linear
IMMEDIATE SOURCE: DNA (genomic)
CLONE: mfd97rs
US-08-222-177A-312

Query Match 0.1%; Score 33; DB 1; Length 33;
Best Local Similarity 100.0%; Pred. No. 0.056;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28944 CTCCTCTCTCACACACACACACACAC 28976
DB 1 CTCCTCTCTCACACACACACACACACAC 33

RESULT 8
US-08-222-177A-175
Sequence 175, Application US/08222177A
Patent No. 5582979
GENERAL INFORMATION:
APPLICANT: Weber, James L.
TITLE OF INVENTION: LENGTH POLYMORPHISMS IN
SEQUENCES AND METHODS OF USING SAME
NUMBER OF SEQUENCES: 460
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dewitt Ross & Stevens, S.C.
STREET: 8000 Excelsior Drive, Suite 401
CITY: Madison
STATE: Wisconsin
COUNTRY: USA
ZIP: 53717-1914
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/222,177A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/341,562
FILING DATE: 21-APR-1989
ATTORNEY/AGENT INFORMATION:
NAME: Sara, Charles S.
REGISTRATION NUMBER: 30,492
REFERENCE/DOCKET NUMBER: 09865,601
TELECOMMUNICATION INFORMATION:
TELEPHONE: (608) 831-2100
TELEFAX: (608) 831-2106
TELEX:
INFORMATION FOR SEQ ID NO: 175:
SEQUENCE CHARACTERISTICS:
LENGTH: 40 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
MOLECULE TYPE: linear
IMMEDIATE SOURCE: DNA (genomic)
CLONE: mfd42rs
US-08-222-177A-175

Query Match 0.1%; Score 33; DB 1; Length 40;
Best Local Similarity 100.0%; Pred. No. 0.054;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28954 CACACACACACACACACACACACACAC 28986
DB 1 CACACACACACACACACACACACACACAC 33

RESULT 9
US-08-222-177A-71
Sequence 71, Application US/08222177A
Patent No. 5582979
GENERAL INFORMATION:
APPLICANT: Weber, James L.
TITLE OF INVENTION: LENGTH POLYMORPHISMS IN
SEQUENCES AND METHODS OF USING SAME
NUMBER OF SEQUENCES: 460
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dewitt Ross & Stevens, S.C.
STREET: 8000 Excelsior Drive, Suite 401
CITY: Madison
STATE: Wisconsin
COUNTRY: USA
ZIP: 53717-1914
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/222,177A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/341,562
FILING DATE: 21-APR-1989
ATTORNEY/AGENT INFORMATION:
NAME: Sara, Charles S.
REGISTRATION NUMBER: 30,492
REFERENCE/DOCKET NUMBER: 09865,601
TELECOMMUNICATION INFORMATION:

TELEPHONE: (608) 831-2100
TELEFAX: (608) 831-2106
TELEX:
INFORMATION FOR SEQ ID NO: 71:
SEQUENCE CHARACTERISTICS:
LENGTH: 46 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-222-177A-71

Query Match 0.1%; Score 33; DB 1; Length 46;
Best Local Similarity 100.0%; Pred. No. 0.052;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28954 CACACACACACACACACACACACACAT 28986
|||||
Db 9 CACACACACACACACACACACACACAT 41

RESULT 10
US-08-222-177A-317
Sequence 317, Application US/08222177A
Patent No. 5582979
GENERAL INFORMATION:
APPLICANT: Weber, James L.
TITLE OF INVENTION: LENGTH POLYMORPHISMS IN
TITLE OF INVENTION: (dc-da)n.(dg-dt)n SEQUENCES AND METHODS OF USING SAME
NUMBER OF SEQUENCES: 460
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dewitt Ross & Stevens, S.C.
STREET: 8000 Excelsior Drive, Suite 401
CITY: Madison
STATE: Wisconsin
COUNTRY: USA
ZIP: 53717-1914
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/222,177A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/341,562
FILING DATE: 21-APR-1989
ATTORNEY/AGENT INFORMATION:
NAME: Sara, Charles S.
REGISTRATION NUMBER: 30,492
REFERENCE/DOCKET NUMBER: 09865.601
TELECOMMUNICATION INFORMATION:
TELEPHONE: (608) 831-2100
TELEFAX: (608) 831-2106
TELEX:
INFORMATION FOR SEQ ID NO: 317:
SEQUENCE CHARACTERISTICS:
LENGTH: 51 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
IMMEDIATE SOURCE:
CLONE: mfg99rs
US-08-222-177A-317

Query Match 0.1%; Score 33; DB 1; Length 51;
Best Local Similarity 100.0%; Pred. No. 0.052;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28953 TCACACACACACACACACACACACACA 28985
|||||
Db 10 TCACACACACACACACACACACACACA 42

RESULT 11
US-08-469-802B-27
Sequence 27, Application US/08469802B
Patent No. 5741645
GENERAL INFORMATION:
APPLICANT: Orr, Harry T.
APPLICANT: Raun, Laura P.W.
APPLICANT: Chung, Ming-Yi
APPLICANT: Zoghbi, Huda Y.
TITLE OF INVENTION: Gene Sequence for Spinocerebellar Ataxia
Patent No. 5741645
TITLE OF INVENTION: Type 1 and Method for Diagnosis
NUMBER OF SEQUENCES: 47
CORRESPONDENCE ADDRESS:
ADDRESSEE: Mueling, Raasch, Gebhardt & Schwappach, P.A.
STREET: 119 No. 5741645th Fourth Street, Suite 203
CITY: Minneapolis
STATE: MN
COUNTRY: USA
ZIP: 55401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/469,802B
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Mueling, Ann M.
REGISTRATION NUMBER: 33,977
REFERENCE/DOCKET NUMBER: 110.00030101
TELECOMMUNICATION INFORMATION:
TELEPHONE: 612-305-1217
TELEFAX: 612-305-1225
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 54 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-469-802B-27

Query Match 0.1%; Score 33; DB 2; Length 54;
Best Local Similarity 100.0%; Pred. No. 0.051;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28954 CACACACACACACACACACACACACAT 28986
|||||
Db 1 CACACACACACACACACACACACACAT 33

RESULT 12
US-08-267-803B-45
Sequence 45, Application US/08267803B
Patent No. 5834183
GENERAL INFORMATION:
APPLICANT: Orr, Harry T.
APPLICANT: Raun, Laura P.W.
APPLICANT: Chung, Ming-Yi
APPLICANT: Zoghbi, Huda Y.
TITLE OF INVENTION: Gene Sequence for Spinocerebellar Ataxia
Patent No. 5834183
TITLE OF INVENTION: Type 1 and Method for Diagnosis

NUMBER OF SEQUENCES: 85
CORRESPONDENCE ADDRESS:
ADDRESSEE: Mueling, Raasch, Gebhardt & Schwappach, P.A.
STREET: P.O. Box 581415
CITY: Minneapolis
STATE: MN
COUNTRY: USA
ZIP: 55458-1415
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/267, 803B
FILING DATE: 28-JUN-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: McCormack, Myra H.
REGISTRATION NUMBER: 36,602
REFERENCE/DOCKET NUMBER: 110,00030120
TELECOMMUNICATION INFORMATION:
TELEPHONE: 612-305-1217
TELEFAX: 612-305-1228
INFORMATION FOR SEQ ID NO: 45:
SEQUENCE CHARACTERISTICS:
LENGTH: 54 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-267-803B-45

Query Match 0.1%; Score 33; DB 3; Length 54;
Best Local Similarity 100.0%; Pred. No. 0.051;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28954 CACACACACACACACACACACACACACACAT 28986
1 CACACACACACACACACACACACACACACAT 33

RESULT 13
5506118-3/C
PATENT NO. 5506118
APPLICANT: BERG, DAVID T.; GRINNELL, BRIAN W.
TITLE OF INVENTION: METHOD OF USING EUKARYOTIC EXPRESSION
VECTORS COMPRISING A POLY GT ELEMENT IN THE PRESENCE OF
TRANS-ACTING GENE PRODUCTS
NUMBER OF SEQUENCES: 13
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/110,475
FILING DATE: 23-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 914,651
FILING DATE: 15-JUL-1992
APPLICATION NUMBER: 255,203
FILING DATE: 07-OCT-1988
SEQ ID NO: 3
LENGTH: 62
5506118-3

Query Match 0.1%; Score 33; DB 7; Length 62;
Best Local Similarity 100.0%; Pred. No. 0.05;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28953 TCACACACACACACACACACACACACACA 28985
53 TCACACACACACACACACACACACACACA 21

RESULT 14
US-08-222-177A-59
Sequence 59, Application US/08222177A
Patent No. 5582979

GENERAL INFORMATION:
APPLICANT: Weber, James L.
TITLE OF INVENTION: LENGTH POLYMORPHISMS IN
(dc-da)n.(dc-dt)n SEQUENCES AND METHODS OF USING SAME
NUMBER OF SEQUENCES: 460
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dewitt Ross & Stevens, S.C.
STREET: 8000 Excelsior Drive, Suite 401
CITY: Madison
STATE: Wisconsin
COUNTRY: USA
ZIP: 53717-1914

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/222,177A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/341,562
FILING DATE: 21-APR-1989

ATTORNEY/AGENT INFORMATION:
NAME: Sara, Charles S.
REGISTRATION NUMBER: 30,492
REFERENCE/DOCKET NUMBER: 09865, 601
TELECOMMUNICATION INFORMATION:
TELEPHONE: (608) 831-2100
TELEFAX: (608) 831-2106

INFORMATION FOR SEQ ID NO: 59:
SEQUENCE CHARACTERISTICS:
LENGTH: 33 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
IMMEDIATE SOURCE:
CLONE: mfd3rs

US-08-222-177A-59

Query Match 0.1%; Score 32; DB 1; Length 33;
Best Local Similarity 100.0%; Pred. No. 0.12;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28954 CACACACACACACACACACACACACACA 28985
1 CACACACACACACACACACACACACACA 32

RESULT 15
US-08-222-177A-218
Sequence 218, Application US/08222177A
Patent No. 5582979

GENERAL INFORMATION:
APPLICANT: Weber, James L.
TITLE OF INVENTION: LENGTH POLYMORPHISMS IN
(dc-da)n.(dc-dt)n SEQUENCES AND METHODS OF USING SAME
NUMBER OF SEQUENCES: 460
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dewitt Ross & Stevens, S.C.
STREET: 8000 Excelsior Drive, Suite 401
CITY: Madison
STATE: Wisconsin
COUNTRY: USA
ZIP: 53717-1914

QY 28953 TCACACACACACACACACACACACACACA 28985
53 TCACACACACACACACACACACACACACA 21

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 31, 2000, 02:52:55 ; Search time 8179.94 Seconds
(without alignments)
13847.741 Million cell updates/sec

Title: US-08-852-495c-1_COPY_165000_195000
Perfect score: 30001
Sequence: 1 AGTGAATTTATATAGAGTC.....GCAGTTTCATTGTCTGTGT 30001

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 4538634 seqs, 1887831982 residues

Word size : 0

Total number of hits satisfying chosen parameters: 130662

Minimum DB seq length: 8
Maximum DB seq length: 105

Post-processing: Listing first 45 summaries

Database :

EST: *
1: em_est1: *
2: em_est2: *
3: em_est3: *
4: em_est4: *
5: em_est5: *
6: em_est6: *
7: em_est7: *
8: em_est8: *
9: em_est9: *
10: em_est10: *
11: em_est11: *
12: em_est12: *
13: em_est13: *
14: em_est14: *
15: em_est15: *
16: em_est16: *
17: em_est17: *
18: em_est18: *
19: em_est19: *
20: gb_est1: *
21: gb_est2: *
22: gb_est3: *
23: gb_est4: *
24: gb_est5: *
25: gb_est6: *
26: gb_est7: *
27: gb_est8: *
28: gb_est9: *
29: gb_est10: *
30: gb_est11: *
31: gb_est12: *
32: gb_est13: *
33: gb_est14: *
34: gb_est15: *
35: gb_est16: *
36: gb_est17: *
37: gb_est18: *
38: gb_est19: *
39: gb_est20: *
40: gb_est21: *
41: gb_est22: *
42: gb_est23: *
43: gb_est24: *

44: gb_est25: *
45: gb_est26: *
46: gb_est27: *
47: gb_est28: *
48: gb_est29: *
49: gb_est30: *
50: gb_est31: *
51: gb_est32: *
52: em_est20: *
53: em_est21: *
54: em_est22: *
55: em_est23: *
56: em_est24: *
57: em_est25: *
58: em_est26: *
59: gb_est33: *
60: gb_est34: *
61: gb_est35: *
62: gb_est36: *
63: gb_est37: *
64: gb_est38: *
65: em_est27: *
66: em_est28: *
67: em_est29: *
68: em_est30: *
69: gb_est39: *
70: gb_est40: *
71: gb_est41: *
72: gb_est42: *
73: gb_est43: *
74: gb_est44: *
75: em_est31: *
76: em_est32: *
77: em_est33: *
78: em_est34: *
79: gb_est45: *
80: gb_est46: *
81: gb_est47: *
82: gb_est48: *
83: em_est35: *
84: em_est36: *
85: em_est37: *
86: em_est38: *
87: gb_est49: *
88: gb_est50: *
89: gb_est51: *
90: gb_est52: *
91: gb_est53: *
92: em_est39: *
93: em_est40: *
94: em_est41: *
95: em_est42: *
96: em_est43: *
97: em_est44: *
98: em_est45: *
99: gb_est54: *
100: gb_est55: *
101: em_est46: *
102: gb_est56: *
103: gb_est57: *
104: gb_est58: *
105: gb_est59: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	45	0.1	75	25 N68192	N68192 za11f07.s1

REFERENCE	AUTHORS	TITLE
1 (pages 1 to 95)	Hallier, L., Allen, M., Bowles, L., Dubbque, T., Geisel, G., Jost, S., Kitzman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J., Moore, B., Scheilenberg, K., Stepcoe, M., Tan, F., Theising, B., White, Y., Wylie, T., Waterson, R. and Wilson, R.	WashU-NCI human EST Project

JOURNAL Unpublished (1997)
COMMENT On Nov 6, 1997 this sequence version replaced g1:932677.
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: estewatson.wustl.edu

This clone is available royalty-free through LNL ; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
Seq primer: -40m13 fwd. ET from Amersham.

FEATURES

source

Location/Qualifiers

1..95
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="IMAGE:856819"
/clone_lib="Stratagene ovary (#937217)"
/sex="female"
/dev_stage="49 year old"
/lab_host="SOLR cells (kanamycin resistant)"
/note="Organ: ovary; Vector: Bluescript SK; Site_1: EcoRI;
Site_2: XhoI; Cloned unidirectionally. Primer: Oligo dt.
Total ovary tissue, normal, caucasian. Average insert
size: 0.8 kb; Uni-ZAP XR Vector: -5' adaptor sequence: 5'
GAATTCGCGACGAG 3' -3' adaptor sequence: 5'
CTCGAGTTTCTTTTCTTTTCTTTT 3' "

BASE COUNT

15 a 30 c 24 g 26 t

ORIGIN

Query Match 0.1%; Score 40; DB 37; Length 95;
Best Local Similarity 100.0%; Pred. No. 0.045; Mismatches 0; Indels 0; Gaps 0;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 25676 TGATCTGCTCACTGCACCTCGCCTCTGCTGCTCAAG 25715
|||||
Db 30 TGATCTGCTCACTGCACCTCGCCTCTGCTGCTCAAG 69

RESULT 8

T63772

LOCUS 105 bp mRNA EST 17-FEB-1995
DEFINITION YC21F01.s1 Stratagene lung (#937210) Homo sapiens cDNA clone
IMAGE:81337 3' similar to contains Alu repetitive element; mRNA
sequence.
T63772 GI:667637

ACCESSION T63772
VERSION T63772.1
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 105)
Hiller, L., Lennon, G., Becker, M., Bonaldo, M.F., Chiappelli, B.,
Chisoe, S., Dietrich, N., Dubuque, T., Favello, A., Gish, W.,
Hawkins, M., Hultman, M., Kucaba, T., Lacy, M., Le, M., Le, N.,
Mardis, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L.,
Rohling, T., Scheinberg, K., Soares, M.B., Tan, F., Thierrey-Meg, J.,
Treviski, E., Underwood, K., Wohldmann, P., Waterston, R., Wilson, R.
and Marra, M.

Generation and analysis of 280,000 human expressed sequence tags
Genome Res. 6 (9), 807-828 (1996)
9704478

TITLE
JOURNAL MEDLINE
COMMENT
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: estewatson.wustl.edu
Insert Size: 513
High quality sequence stops: 69 source: IMAGE Consortium, LNL This
clone is available royalty-free through LNL ; contact the IMAGE
Consortium (info@image.lnl.gov) for further information.

Insert Length: 513 Std Error: 0.00
Seq primer: -21m13
High quality sequence stop: 69.
Location/Qualifiers
1..105

FEATURES

source

/organism="Homo sapiens"
/db_xref="GDB:484954"
/db_xref="taxon:9606"
/clone_image="IMAGE:81337"
/clone_lib="Stratagene lung (#937210)"
/sex="male"
/dev_stage="72 years"
/lab_host="SOLR cells (kanamycin resistant)"
/note="Organ: lung; Vector: pBluescript SK; Site_1:
EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer:
Oligo dt. normal lung. Average insert size: 1.0 kb;
Uni-ZAP XR Vector: -5' adaptor sequence: 5' GAATTCGCGACGAG
3' -3' adaptor sequence: 5' CTCGAGTTTCTTTTCTTTTCTTTT 3' "

BASE COUNT

17 a 34 c 29 g 25 t

ORIGIN

Query Match 0.1%; Score 40; DB 21; Length 105;
Best Local Similarity 100.0%; Pred. No. 0.043; Mismatches 0; Indels 0; Gaps 0;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3372 ACCTCGCTCTCCGCGGTTCAAGGATTCCTGCTCAGC 3411
|||||
Db 36 ACCTCGCTCTCCGCGGTTCAAGGATTCCTGCTCAGC 75

RESULT 9

AA916300 92 bp mRNA EST 17-JUN-1998
LOCUS on22c08.s1 NCI-CGAP Lu5 Homo sapiens cDNA clone IMAGE:1557422 3'
DEFINITION similar to contains Alu repetitive element; mRNA sequence.
AA916300
ACCESSION AA916300.1 GI:3055692

VERSION AA916300.1
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 92)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
On May 5, 1995 this sequence version replaced g1:797742.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
DNA Sequencing: Greg Lennon, Ph.D.
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
www.bio.lnl.gov/dbp/image/image.html

Insert Length: 1519 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 61.
Location/Qualifiers
1..92
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="IMAGE:1557422"
/clone_lib="NCI-CGAP Lu5"
/tissue_type="carcinoid"
/lab_host="DH10B"
/note="Organ: lung; Vector: pT713D-Pac (Pharmacia) with a

FEATURES

source

modified polylinker; 1st strand cDNA was prepared from neuroendocrine lung carcinoid, and was then primed with a Not I - oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia) digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library is normalized. Library was constructed by Bento Soares and M. Fatima Bonaldo. "

BASE COUNT 15 a 34 c 21 g 22 t
 ORIGIN
 Query Match 0.1%; Score 39; DB 40; Length 92;
 Best Local Similarity 100.0%; Pred. No. 0.092;
 Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 18014 CACCTCCGGGTTCAGCAATCTCTGCTCAGCCTCC 18052
 DB 45 CACCTCCGGGTTCAGCAATCTCTGCTCAGCCTCC 83

RESULT 10
 LOCUS AA583697/c 101 bp mRNA EST 26-SEP-1997
 DEFINITION nm38f10.s1 NCI-CGAP_K16 Homo sapiens cDNA clone IMAGE:1088107 3'
 similar to contains Alu repetitive element; mRNA sequence.
 AA583697
 ACCESSION AA583697.1 GI:2368306
 VERSION EST.
 KEYWORDS human.
 SOURCE Homo sapiens
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE
 AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT On Apr 14, 1993 this sequence version replaced gi:692709.
 Contact: Robert Strausberg, Ph.D.
 Tel: (301) 496-1550
 Email: Robert.Strausberg@nih.gov
 Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R.
 Emmert-Buck, M.D., Ph.D.
 CDNA Library Preparation: Stratagene, Inc.
 CDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/BLN at:
 www.bio.linn.gov/bbrp/image/image.html

Insert Length: 1890 Std Error: 0.00
 Seg primer: -40m13 fwd. ET from Amersham
 High quality sequence stop: 93.
 Location/Qualifiers

FEATURES
 Source 1..101
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:1088107"
 /clone_id="NCI-CGAP_K16"
 /sex="mixed"
 /tissue_type="kidney tumor"
 /lab_host="SOLR (kanamycin resistant)"
 /note="Organ: kidney; Vector: Bluescript SK-; Site: 1;
 Ecor: Site: 2; Xho: Cloned unidirectionally. Primer:
 oligo dt. Pooled kidney tumors. 5' adaptor sequence: 5'
 GAATTCGGCAGAG 3' 3' adaptor sequence: 5'
 CTCGAGTTTTTTTTTTTTTTT 3' Average insert size: 1.0 kb."

BASE COUNT 21 a 29 c 28 g 23 t
 ORIGIN

Query Match 0.1%; Score 39; DB 35; Length 101;
 Best Local Similarity 100.0%; Pred. No. 0.087;

Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 19733 TCAGCGCTAATCCAGCACTTGGGAGGCCAGGCGG 19771
 DB 101 TCAGCGCTAATCCAGCACTTGGGAGGCCAGGCGG 63

RESULT 11
 LOCUS N98204/c 94 bp mRNA EST 18-NOV-1996
 DEFINITION 0288C3 czappd2.1, Debopam Chakrabarti Plasmodium falciparum cDNA
 clone PF0288C, mRNA sequence.
 ACCESSION N98204
 VERSION N98204.1 GI:1674570
 KEYWORDS EST.
 SOURCE malaria parasite P. falciparum.
 ORGANISM Plasmodium falciparum
 Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.

REFERENCE
 AUTHORS Dame, J.B., Arnott, D.E., Bourke, P., Chakrabarti, D., Christodoulou, Z.,
 1 (bases 1 to 94)
 Coppel, R., Cowman, A., Craig, A., Fischer, K., Foster, J., Goodman, N.,
 Hilderberg, K., Holder, A.A., Holt, D., Kemp, D., Lanzer, M., Lim, A.,
 Newbold, C., Ravetch, J.V., Reddy, G.R., Rubio, J., Schuster, S.M.,
 Su, X.-Z., Thompson, J.K., Vital, F., Wellens, T.E. and Werner, E.
 Current status of the Plasmodium falciparum genome project
 Mol. Biochem. Parasitol. 79, 1-12 (1996)
 97001675
 JOURNAL On May 5, 1995 this sequence version replaced gi:797974.
 COMMENT Contact: Debopam Chakrabarti
 Department of Molecular Biology and Microbiology
 University of Central Florida
 Orlando, FL 32816-2360
 Tel: 407 384 2061
 Fax: 407 384 3095
 Email: dchakr@pegasus.cc.ucf.edu
 Seq primer: 73.
 Location/Qualifiers

FEATURES
 source 1..94
 /organism="Plasmodium falciparum"
 /strain="Dd2"
 /db_xref="taxon:5833"
 /clone="PF0288C"
 /clone_id="czappd2.1, Debopam Chakrabarti"
 /lab_host="E. coli XL-1 blue"
 /note="Vector: Lambda ZAP II; Site: 1; Ecor I; Site: 2; Xho
 I; POLYA+ RNA. from asynchronous blood stage parasites of
 the Dd2 isolate cultured in vitro. was reverse transcribed
 using an oligo dt-Xho I primer. Second strand was
 prepared using RNase H and DNA polymerase I. Ecor I
 adapters were ligated to the cDNA, and it was digested
 with Xho I. Prepared fragments were ligated into Ecor I +
 Xho I digested lambda ZAP II vector. "

BASE COUNT 49 a 4 c 6 g 31 t 4 others
 ORIGIN

Query Match 0.1%; Score 38; DB 25; Length 94;
 Best Local Similarity 100.0%; Pred. No. 0.18;
 Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 21527 ATTATATTATTATTATTATTATTATTATTATTATT 21564
 DB 64 ATTATATTATTATTATTATTATTATTATTATTATT 27

RESULT 12
 LOCUS C05051/c 104 bp mRNA EST 30-JUL-1996
 DEFINITION C05051 Human heart CDNA (YNakamura) Homo sapiens cDNA clone
 3NC4544, mRNA sequence.
 ACCESSION C05051
 VERSION C05051.1 GI:1468302
 KEYWORDS EST.

SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS 1 (bases 1 to 104)
Tanaka,T., Ogiwara,A., Uchiyama,I., Takagi,T., Yazaki,Y. and Nakamura,Y.
TITLE Construction of a normalized directionally cloned cDNA library from adult heart and analysis of 3040 clones by partial sequencing
JOURNAL Genomics 35 (1), 231-235 (1996)
MEDLINE 96299762
COMMENT Contact: Yusuke Nakamura
Institute of Medical Science
University of Tokyo
4-6-1, Shirokanedai, Minato-ku, Tokyo 108, Japan
Tel: 81-3-5449-5372
Fax: 81-3-5449-5433
Email: yusuke@ims.u-tokyo.ac.jp.

FEATURES
source
1..104
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_1lb="3NHc4544"
/clone_1lb="Human heart cDNA (YNakamura)"
/dev_stage="adult"
/note="Organ: heart; normalized directionally cloned cDNA from adult heart"

BASE COUNT 24 a 26 c 37 g 17 t

ORIGIN

Query Match 0.1%; Score 38; DB 27; Length 104;
Best Local Similarity 100.0%; Pred. No. 0.17; 0; Indels 0; Gaps 0;
Matches 38; Conservative 0; Mismatches 0;

Qy 3378 GCCCTCCGGGTTCAAGCATCTCTCGCTCAGCCTCC 3415
|||||
Db 38 GCCTCCGGGTTCAAGCATCTCTCTCGCTCAGCCTCC 1

RESULT 13
AA548135 92 bp mRNA EST 14-AUG-1997
LOCUS AK55611.s1 NCI_CGAP_P7 Homo sapiens cDNA clone IMAGE:1017404, mRNA
DEFINITION Sequence.
ACCESSION AA548135
VERSION AA548135.1 GI:2318417
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 92)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT On Sep 12, 1996 this sequence version replaced gi:1397885.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: David G. Bostwick, M.D., Rodrigo F. Chuang, M.D., Michael R. Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: David B. Kitzman, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
www-bio.lnl.gov/biopr/image/image.html

FEATURES
source
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 70.
Location/Qualifiers

SOURCE 1..92
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_1lb="IMAGE:1017404"
/clone_1lb="NCI_CGAP_P7"
/sex="male"
/tissue_type="low-grade prostatic neoplasia"
/lab_host="DH10B"
/note="Organ: prostate; Vector: pAMP10; mRNA made from prostate intraepithelial neoplasia (high-grade), cDNA made by oligo-dT priming. Non-directionally cloned. Size selected on agarose gel, average insert size 600 bp."

BASE COUNT 30 a 18 c 27 g 17 t

ORIGIN

Query Match 0.1%; Score 37; DB 35; Length 92;
Best Local Similarity 100.0%; Pred. No. 0.37;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18257 TTTTGTGAGACAGAGCTCAGTCTGTACCCAGCT 18293
|||||
Db 86 TTTTGTGAGACAGAGCTCAGTCTGTACCCAGCT 50

RESULT 14
B80126/c 102 bp DNA GSS 24-OCT-1998
LOCUS CIT-HSP-2042J24.TR CIT-HSP Homo sapiens genomic clone 2042J24,
DEFINITION genomic survey sequence.
ACCESSION B80126
VERSION B80126.1 GI:2867149
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 102)
Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K., Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M. and Venter,J.C.
Use of a random BAC End Sequence Database for Sequence-Ready Map Building (1998)
JOURNAL Unpublished (1998)
COMMENT Other_GSSs: CIT-HSP-2042J24.TF
Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tifg.org
Clones are available from Research Genetics (info@resgen.com). BAC end search page:
http://www.tifg.org/tfdb/humgen/bac_end_search/bac_end_search.html
Seq primer: M13 Reverse
Class: BAC ends.

FEATURES
source
1..102
/organism="Homo sapiens"
/db_xref="GDB:705355"
/db_xref="taxon:9606"
/clone_1lb="2042J24"
/clone_1lb="CIT-HSP"
/sex="Male"
/cell_type="Sperm"
/note="Vector: pBelBAC11; Site_1: HindIII; Site_2: HindIII"

BASE COUNT 20 a 32 c 19 g 31 t

ORIGIN

Query Match 0.1%; Score 37; DB 81; Length 102;
 Best Local Similarity 100.0%; Pred. No. 0.35;
 Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Search completed: March 31, 2000, 13:42:06
 Job time: 210380 sec

QY 21782 AGCCTCCCAAGTGTGGATTACAGCGGTGAGCCAC 21818
 |||
 Db 102 AGCCTCCCAAGTGTGGATTACAGCGGTGAGCCAC 66

RESULT 15

T94466 105 bp mRNA EST 24-MAR-1995
 LOCUS Y635D02.1 Striatogene lung (#937210) Homo sapiens cDNA clone
 DEFINITION IMAGE:119691 5' similar to contains Alu repetitive element; mRNA
 sequence.

ACCESSION T94466
 VERSION T94466.1 GI:727954
 KEYWORDS EST.

SOURCE human.
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 105)
 AUTHORS Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chapell,B.,
 Chisoe,S., Dietrich,N., Dubuque,T., Favello,A., Gish,W.,
 Hawkins,M., Hultman,M., Kucaba,T., Lacy,M., Le,N.,
 Mardis,E., Moore,B., Morris,M., Parsons,J., Prange,C., Rifkin,L.,
 Rohlfing,T., Schellendberg,K., Soares,M.B., Tan,F., Thierry-Mieg,J.,
 Trevaaskis,E., Underwood,K., Wohlman,P., Waterston,R., Wilson,R.
 and Marra,M.

TITLE Generation and analysis of 280,000 human expressed sequence tags
 JOURNAL Genome Res. 6 (9), 807-828 (1996)
 MEDLINE 97044478

COMMENT Contact: Wilson RK
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu

Insert Size: 614
 High quality sequence stops: 70 Source: IMAGE Consortium, LIND. This
 clone is available royalty-free through LIND; contact the IMAGE
 Consortium (info@image.llnl.gov) for further information.
 Insert length: 614 Std Error: 0.00
 Seq primer: M13RP1
 High quality sequence stop: 70.
 Location/Qualifiers

FEATURES

source 1..105
 /organism="Homo sapiens"
 /db_xref="GDB:487980"
 /db_xref="taxon:9606"
 /clone="IMAGE:119691"
 /clone_lib="Striatogene lung (#937210)"
 /sex="male"
 /dev_stage="72 years"
 /lab_host="SOLR cells (kanamycin resistant)"
 /note="organ: lung; Vector: pBluescript SK-; Site_1:
 EcorI; Site_2: XhoI; Cloned unidirectionally. Primer:
 Oligo dt. normal lung. Average insert size: 1.0 kb;
 Uni-ZAP XR Vector; -5' adaptor sequence: 5' GAATTCGACGAG
 3' -3' adaptor sequence: 5' CTCGAGTTT TTT TTT TTT 3'."
 BASE COUNT 31 a 20 c 34 g 18 t 2 others
 ORIGIN

Query Match 0.1%; Score 37; DB 21; Length 105;
 Best Local Similarity 100.0%; Pred. No. 0.34;
 Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19898 CTGAGCGAGAGATCACTTGAACCGAGAGCGAGAG 19934
 |||
 Db 57 CTGAGCGAGAGATCACTTGAACCGAGAGCGAGAG 93

Mon Apr 3 08:24:12 2000

us-08-852-495c-1_copy_165000_195000.rst

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 31, 2000, 17:25:40 ; Search time 13412.4 Seconds

(without alignments) 1 updates/sec

Title: US-08-852-495C-1_COPY_190000_220000

Perfect score: 30001
Sequence: 1 GATATATCTCATGTGCTCTG.....AGGGCCAAAGAGATCATG 30001

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 82193 seqs, -1518192014 residues

Word size : 0

Total number of hits satisfying chosen parameters: 345848

Minimum DB seq length: 8

Maximum DB seq length: 105

Post-processing: Listing first 45 summaries

Database : GenEmbl:*

1: gb_ba1:*
2: gb_ba2:*
3: gb_om:*
4: gb_ov:*
5: gb_pal:*
6: gb_ph:*
7: gb_pl1:*
8: gb_pl2:*
9: gb_pl3:*
10: gb_pr2:*
11: gb_pr3:*
12: gb_ro:*
13: gb_sts:*
14: gb_sy:*
15: gb_un:*
16: gb_vl:*
17: em_fun:*
18: em_hum1:*
19: em_hum2:*
20: em_in:*
21: em_om:*
22: em_or:*
23: em_ov:*
24: em_pat:*
25: em_ph:*
26: em_pl:*
27: em_ro:*
28: em_sts:*
29: em_sy:*
30: em_un:*
31: em_vl:*
32: gb_htg1:*
33: gb_htg2:*
34: gb_in1:*
35: gb_in2:*
36: em_ba1:*
37: em_ba2:*
38: em_hum3:*
39: em_hum4:*
40: gb_pr4:*
41: gb_htg3:*
42: gb_htg4:*
43: gb_htg5:*

44: gb_htg6:*
45: gb_htg7:*
46: em_htg1:*
47: em_htg2:*
48: em_htg3:*
49: em_hum5:*
50: gb_pl3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	54	0.2	60	5 I31332	I31332 Sequence 24
2	50	0.2	91	5 I31254	I31254 Sequence 16
3	50	0.2	92	5 I31518	I31518 Sequence 43
4	46	0.2	66	40 AF087511	AF087511 Homo sapi
5	43	0.1	65	9 HUMALUNCB	L36836 Homo sapien
6	42	0.1	56	5 I31153	I31153 Sequence 65
7	41	0.1	50	5 I31467	I31467 Sequence 37
8	40	0.1	40	5 A68621	A68621 Sequence 1
9	40	0.1	85	5 A45374	A45374 Sequence 44
10	40	0.1	85	5 AR061179	AR061179 Sequence
11	40	0.1	85	5 HSMG3B103	X86915 H. sapiens s
12	37	0.1	60	5 AR051487	AR051487 Sequence
13	37	0.1	76	5 AR051499	AR051499 Sequence
14	36	0.1	80	9 HOMBRKFAC	M36133 Human alpha
15	36	0.1	101	10 S79560	S79560 HRX (intron
16	36	0.1	101	10 S79561	S79561 dHRX (partl
17	35	0.1	33	5 A25212	A25212 Inter-Alu s
18	35	0.1	35	24 E09140	E09140 Synthetic D
19	35	0.1	60	11 HSCBFB5108	AF084962 Homo sapi
20	35	0.1	68	13 HUMUT997A	L39153 Human STS U
21	35	0.1	83	13 G37817	G37817 SAH Plasmod
22	35	0.1	85	5 AR051522	AR051522 Sequence
23	35	0.1	90	9 HUMLDLRF1	K03555 Human low d
24	34	0.1	41	9 HUMALUNCA	L36835 Homo sapien
25	34	0.1	67	9 HUMALUNCA	L36843 Homo sapien
26	34	0.1	72	13 G31952	G31952 MCM92 Chick
27	34	0.1	75	3 BRIG445	X85061 B. taurus co
28	34	0.1	88	3 OAU15702	U15702 Ovis aries
29	34	0.1	95	13 AU026531	AU026531 Rattus no
30	34	0.1	95	13 AU026788	AU026788 Rattus no
31	34	0.1	99	13 G37948	G37948 C4M78 Plasm
32	34	0.1	99	13 G37948	G37948 C4M78 Plasm
33	34	0.1	101	13 G37906	G37906 C3M4 Plasm
34	33	0.1	33	5 I31400	I31400 Sequence 31
35	33	0.1	40	5 I31263	I31263 Sequence 17
36	33	0.1	46	5 I31159	I31159 Sequence 71
37	33	0.1	51	5 I31405	I31405 Sequence 31
38	33	0.1	54	5 AR002288	AR002288 Sequence
39	33	0.1	54	5 AR053139	AR053139 Sequence
40	33	0.1	60	34 DBOR81	M36554 D. melanogas
41	33	0.1	62	5 I07633	I07633 Sequence 2
42	33	0.1	79	10 S73203	S73203 AL1-1 (land
43	33	0.1	82	3 HUMRSCB	K02284 Human poly
44	33	0.1	83	13 G37817	G37817 SAH Plasmod
45	33	0.1	87	9 HUM1	L23536 Human Chrom

ALIGNMENTS

RESULT 1

I31332 I31332 60 bp DNA

LOCUS I31332

DEFINITION Sequence 244 from patent US 5582979.

ACCESSION I31332

VERSION I31332.1 GI:1822123

PAT 06-FEB-1997

KEYWORDS	Unknown.
SOURCE	Unknown.
ORGANISM	Unclassified.
REFERENCE	1 (bases 1 to 60)
AUTHORS	Weber, J. L.
TITLE	Length polymorphisms in (dc-da).sub.n.(dg-dt).sub.n sequences and method of using the same
JOURNAL	Patent: US 5582979-A 244 10-DEC-1996;
FEATURES	Location/Qualifiers
source	1..60
BASE COUNT	18 a 30 c 0 g 12 t
ORIGIN	/organism="unknown"

Query Match Score 54; DB 5; Length 60;
Best Local Similarity 100.0%; Pred. No. 1,38-15;
Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3933 CTTCTGTCGTCTGTCATCAACAACAAACAACAACAACAACA 3985
DB 2 CTCTCTCTCTCTCTCTCTCTCAACAACAACAACAACAACAACA 55

RESULT	2				
LOCUS	I31254				
DEFINITION	I31254	91 bp	DNA	PAT	06-FEB-1997
ACCESSION	I31254	Sequence 166 from patent US 5582979.			
VERSION	I31254.1	GI:1822045			
KEYWORDS	.				
SOURCE	Unknown.				
ORGANISM	Unknown.				
REFERENCE	Unclassified.				
AUTHORS	1 (bases 1 to 91)				
TITLE	Weber,J.L.				
JOURNAL	Length polymorphisms in (dC-dA).sub.n.(dG-dT).sub.n sequences and method of using the same Patent: US 5582979-A 166 10-DEC-1996;				
FEATURES	Location/Qualifiers				
source	1..91				
	/organism="unknown"				
BASE COUNT	20 a	43 c	1 g	27 t	
ORIGIN					

Query Match Score 50; DB 5; Length 91;
Best Local Similarity 100.0%; Pred. No. 1-3e-13;
Matches 50; Conservative 0; Mismatches 0; Indels 0;
Gaps 0;

QY 3932 CTTCTCTCTCTCTCTCTCTCTCACAACACACACACACACAGA 3981
|||||
Db 30 TTTTCTCTCTCTCTCTCTCTCTCACACACACACACACACACAGCA 79

RESULT	3		
LOCUS	131518		
DEFINITION	Sequence 430 from patent US 5582979.	PAT	06-FEB-1997
ACCESSION	131518		
VERSION	131518.1	GI:1822309	
KEYWORDS	.		
SOURCE	Unknown.		
ORGANISM	Unclassified.		
REFERENCE	1 (bases 1 to 92)		
AUTHORS	Weber, J.L.		
TITLE	Length polymorphisms in (dC-dA).sub.n.(dG-dT).sub.n sequences and method of using the same		
JOURNAL	Patent: US 5582979-A 430 10-DEC-1996;		
FEATURES	Location/Qualifiers		
source	1..92		

		/organism="unknown"			
BASE COUNT	21 a	43 c	1 g	27 t	
ORIGIN					
Query Match		0.2%;	Score 50;	DB 5;	Length 92:
Best Local Similarity		100.0%;	Pred. No.	1.3e-13;	
Matches	50; Conservative	0;	Mismatches	0;	Indels
				0;	Gaps
					0.
OY	3932 CTCCTCCTCTCTCTCTCTCTCTCACACACACACACACACACA				3981
Db	30 CTCCTCTCTCTCTCTCTCTCTCTCACACACACACACACACACA				79

AF087511/c 4 RESULT 13-SEP-1999

LOCUS 66 bp mRNA PRI

DEFINITION Homo sapiens clone ENac+22 epithelial sodium channel alpha subunit (SCNN1A) mRNA, alternatively spliced, partial sequence.

ACCESSION AF087511.1 GI:5870626

VERSION 1

KEYWORDS human.

SOURCE Homo sapiens

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Euthera; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 66)

AUTHORS Oh.Y. and Warnock,D.G.

TITLE An Alu cassette in the human epithelial sodium channel

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 66)

AUTHORS Oh.Y. and Warnock,D.G.

TITLE Direct Submission

JOURNAL Submitted (26-AUG-1998) Medicine, UAB, 1720 7th Ave. SO., Birmingham, AL 35294, USA

FEATURES	Location/Qualifiers
source	1..66 /organism="Homo sapiens" /db_xref="taxon:9606" /chromosome="12" /clone="ENAC+22"
misc_feature	<1..>66 /gene="SCNN1A" /note="alternatively spliced epithelial sodium channel alpha subunit containing intronic Alu sequence"
gene	<1..>66 /gene="SCNN1A"
BASE COUNT	12 a 20 c 24 g 10 t
ORIGIN	

Query Match	Score 46;	DB 40;	Length 66;
Best Local Similarity	100.0%;	Pred. No. 1.3e-11;	
Matches 46;	Conservative 0;	Mismatches 0;	Indels 0;
Gaps 0;			
QY 24738	CTCTGATATCCGCGCCGCTCGCCTCCCAAGATGCTGGATTCAGG	24783	
Db 66	CTCTGATATCCGCGCCGCTCGCCTCCCAAGATGCTGGATTCAGG	21	

	RESULT	5	PRI	08-OCT-1994
HUMALUANC/c				
LOCUS	HUMALUANC	65 bp	DNA	
DEFINITION	Homo sapiens 4000 year old remains from Nekht-anhk Alu repeat 12:2			
ACCESSION	L36836			
VERSION	L36836.1	GI:556194		
KEYWORDS				
SOURCE	Homo sapiens (individual_isolate 4000 year old remains from Nekht-anhk) liver DNA.			
ORGANISM	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;			
	Euteria; Primates; Catarrhini; Hominoideae; Homo.			
REFERENCE	1 (sites)			

AUTHORS Paabo,S.
TITLE Ancient DNA: extraction, characterization, molecular cloning, and enzymatic amplification
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 86 (6), 1939-1943 (1989)
MEDLINE 89184542

FEATURES
source location/Qualifiers
1. .65

/organism="Homo sapiens"
/isolate="4000 year old remains from Nekht-anh"
/db_xref="taxon:9606"
/tissue_type="liver"
repeat_region <1. .>65
/rpt_family="Alu"

BASE COUNT 17 a 17 c 21 g 10 t
ORIGIN

Query Match 0.1%; Score 43; DB 9; Length 65;
Best Local Similarity 100.0%; Pred. No. 4.2e-10;
Matches 43; Conservative 0; Mismatches 0; Indels 0;

OY 24735 GACCTCGATCCGCCGCTCCGCTCCCAAGTCTGGAT 24777
DB 65 GACCTCGATCCGCCGCTCCGCTCCCAAGTCTGGAT 23

RESULT 6
LOCUS I31153 56 bp DNA PAT 06-FEB-1997
DEFINITION Sequence 65 from patent US 5582979.
ACCESSION I31153
VERSION I31153.1 GI:1821944
KEYWORDS
SOURCE Unknown.
ORGANISM Unidentified.

REFERENCE 1 (bases 1 to 56)

AUTHORS Weber,J.L.
TITLE Length polymorphisms in (dc-da).sub.n.(dg-dt).sub.n sequences and method of using the same

JOURNAL Patent: US 5582979-A 65 10-DEC-1996;
FEATURES
source location/Qualifiers
1. .56

BASE COUNT 23 a 28 c 0 g 5 t
ORIGIN

Query Match 0.1%; Score 42; DB 5; Length 56;
Best Local Similarity 100.0%; Pred. No. 1.3e-09;
Matches 42; Conservative 0; Mismatches 0; Indels 0;

OY 3944 CTCCTCTCTCACACACACACACACACACACACA 3985
DB 1 CTCCTCTCTCACACACACACACACACACACACACACA 42

RESULT 7
LOCUS I31467 50 bp DNA PAT 06-FEB-1997
DEFINITION Sequence 379 from patent US 5582979.
ACCESSION I31467
VERSION I31467.1 GI:1822258
KEYWORDS
SOURCE Unknown.
ORGANISM Unidentified.

REFERENCE 1 (bases 1 to 50)

AUTHORS Weber,J.L.
TITLE Length polymorphisms in (dc-da).sub.n.(dg-dt).sub.n sequences and method of using the same

JOURNAL Patent: US 5582979-A 379 10-DEC-1996;
FEATURES
source location/Qualifiers

source 1. .50
BASE COUNT 20 a 25 c 0 g 5 t
ORIGIN

Query Match 0.1%; Score 41; DB 5; Length 50;
Best Local Similarity 100.0%; Pred. No. 4.3e-09;
Matches 41; Conservative 0; Mismatches 0; Indels 0;

OY 3945 TTCCTCTCTCACACACACACACACACACACACA 3985
DB 1 TTCCTCTCTCACACACACACACACACACACACACA 41

RESULT 8
LOCUS A68621/c 40 bp DNA PAT 06-MAY-1999
DEFINITION Sequence 1 from Patent WO9801573.
ACCESSION A68621
VERSION A68621.1 GI:4759648
KEYWORDS
SOURCE Unidentified.
ORGANISM Unidentified.

REFERENCE 1 (bases 1 to 40)

AUTHORS Resnick M.A., Laktionov V.L., Koupina N.Y. and Perkins E.L.
TITLE TRANSFORMATION-ASSOCIATED RECOMBINATION CLONING

JOURNAL Patent: WO 9801573-A 15-JAN-1998;
US HEALTH (US)

FEATURES
source location/Qualifiers
1. .40
/organism="unidentified"
/db_xref="taxon:32644"

BASE COUNT 7 a 12 c 13 g 8 t
ORIGIN

Query Match 0.1%; Score 40; DB 5; Length 40;
Best Local Similarity 100.0%; Pred. No. 1.4e-08;
Matches 40; Conservative 0; Mismatches 0; Indels 0;

OY 24755 TCGGCTCCCAAGTCTGGGATTCAGCGGTGAGCCACC 24794
DB 40 TCGGCTCCCAAGTCTGGGATTCAGCGGTGAGCCACC 1

RESULT 9
LOCUS A45374/c 85 bp DNA PAT 07-MAR-1997
DEFINITION Sequence 44 from Patent WO9517522.
ACCESSION A45374
VERSION A45374.1 GI:2299846
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.

REFERENCE 1 (bases 1 to 85)

AUTHORS Jeffreys,A.J. and Armour,J.
TITLE IDENTIFICATION OF SIMPLE TANDEM REPEATS

JOURNAL Patent: WO 9517522-A 44 29-JUN-1995;
UNIV LEICESTER (GB)

COMMENT Other publication AU 1277995 950710.
FEATURES
source location/Qualifiers
1. .85

BASE COUNT 58 a 1 c 24 g 2 t
ORIGIN

Query Match 0.1%; Score 40; DB 5; Length 85;
Best Local Similarity 100.0%; Pred. No. 1.3e-08;

KEYWORDS Fabry disease; breakpoint junction; glycosphingolipid catabolism.
SOURCE Human DNA.
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
AUTHORS Eutheria; Primates; Catarrhini; Homiidae; Homo.
TITLE 1 (bases 1 to 80)
Kornreich, R., Bishop, D.F. and Desnick, R.J.
Alpha-galactosidase A gene rearrangements causing Fabry disease:
identification of short direct repeats at breakpoints in an
Alu-rich gene
J. Biol. Chem. 265, 9319-9326 (1990)

JOURNAL MEDLINE
FEATURES location/Qualifiers
source 1. 80
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 18 a 28 c 17 g 17 t
ORIGIN

Query Match 0.1%; Score 36; DB 9; Length 80;
Best Local Similarity 100.0%; Pred. No. 1.4e-06;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18946 GCCTCGGCTCCCAAGTCTGGGATTACAGGATG 18981
|||||
Db 16 GCCTCGGCTCCCAAGTCTGGGATTACAGGATG 51

RESULT 15
S79560 579560 101 bp DNA PRI 27-JAN-1996
LOCUS HRX [intron 6] [human, Genomic, 101 nt].
DEFINITION S79560
ACCESSION S79560
VERSION S79560.1 GI:1168041
KEYWORDS human.
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
REFERENCE Eutheria; Primates; Catarrhini; Homiidae; Homo.
AUTHORS 1 (bases 1 to 101)
Bernard, O.A., Romana, S.P., Schichman, S.A., Mauchauffe, M.,
Jonveaux, P. and Berger, R.
Partial duplication of HRX in acute leukemia with trisomy 11
Leukemia 9 (9), 1487-1490 (1995)
95387660
JOURNAL MEDLINE
REMARK GenBank staff at the National Library of Medicine created this
entry [NCBI g1bbsq 170933] from the original journal article.
This sequence comes from Fig. 3.
Map location: 11q23.

FEATURES location/Qualifiers
source 1. 101
/organism="Homo sapiens"
/db_xref="taxon:9606"

gene 1. 101
/partial
/gene="HRX"

BASE COUNT 27 a 21 c 28 g 25 t
ORIGIN

Query Match 0.1%; Score 36; DB 10; Length 101;
Best Local Similarity 100.0%; Pred. No. 1.3e-06;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14193 GAGACGAGCTGGCCACATGTGTAACCTGTCTC 14228
|||||
Db 11 GAGACGAGCTGGCCACATGTGTAACCTGTCTC 46

KW linkage analysis; genetic disease; animal; plant; breeding; locus;
KM hybridisation; chromosome; ds.
OS Homo sapiens.
PN US5582979-A.
PD 10-DEC-1996.
PF 21-APR-1989; 341562.
PR 21-APR-1989; US-341562.
PR 05-SEP-1991; US-754351.
PR 04-APR-1994; US-222177.
PA (MARS-) MARSHFIELD CLINIC.
PI Weber JL.
DR WPI; 97-042299/04.
PT Detection of polymorphic genetic markers of the form
PT (dc-da)n(dg-dt)n - using novel nucleic acid mols. as primers
PS Disclosure: Column 9-10; 186pp; English.
CC The invention relates to the isolation of polymorphic repeat sequences
CC having the sequence (dc-da)n.(dg-dt)n which can be used as genetic
CC markers. Primers based on these sequences can be used to detect these
CC repeats, especially for use in e.g. paternity or maternity testing,
CC human genetic analysis such as linkage analysis of genetic disease,
CC commercial animal or plant breeding or pedigree analysis. Clones
CC containing the repeat sequences were isolated by hybridisation of
CC chromosome-specific phage libraries with a synthetic poly(dc-da).(dg-dt)
CC probe. Over 100 repeat blocks were isolated. The inserts from the clones
CC were amplified by primers T65798-T66047. Those clones where the repeat
CC sequence has been determined are shown in T65704-797. This repeat
CC sequence is from the marker clone Md39 which contains the repeat
CC sequence having the formula: (TC)12.5GTT(TC)11.5(CA)14A(CA)5.5.
SQ Sequence 91 BP; 20 A; 43 C; 1 G; 27 T;

Query Match 0.2%; Score 50; DB 1; Length 91;
Best Local Similarity 100.0%; Pred. No. 2.9e-07;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3332 CTCCTCTCTCTCTCTCTCTCTCTCACACACACACACACACACACA 3981
Db 30 CTCCTCTCTCTCTCTCTCTCTCTCACACACACACACACACACACA 79

RESULT 3
T66081
ID T66081 standard; DNA; 92 BP.
AC T66081;
DT 18-JUN-1997 (first entry)
DE (dc-da)n.(dg-dt)n polymorphic repeat sequence #12.
KM Polymorphism; repeat sequence; genetic marker; primer; amplification;
KM PCR; polymerase chain reaction; paternity; maternity; human; pedigree;
KM linkage analysis; genetic disease; animal; plant; breeding; locus;
KM hybridisation; chromosome; ds.
OS Homo sapiens.
PN US5582979-A.
PD 10-DEC-1996.
PF 21-APR-1989; 341562.
PR 21-APR-1989; US-341562.
PR 05-SEP-1991; US-754351.
PR 04-APR-1994; US-222177.
PA (MARS-) MARSHFIELD CLINIC.
PI Weber JL.
DR WPI; 97-042299/04.
PT Detection of polymorphic genetic markers of the form
PT (dc-da)n(dg-dt)n - using novel nucleic acid mols. as primers
PS Example 8; Column 57-58; 186pp; English.
CC The invention relates to the isolation of polymorphic repeat sequences
CC having the sequence (dc-da)n.(dg-dt)n which can be used as genetic
CC markers. Primers based on these sequences can be used to detect these
CC repeats, especially for use in e.g. paternity or maternity testing,
CC human genetic analysis such as linkage analysis of genetic disease,
CC commercial animal or plant breeding or pedigree analysis. Clones
CC containing the repeat sequences were isolated by hybridisation of
CC chromosome-specific phage libraries with a synthetic poly(dc-da).(dg-dt)
CC probe. Over 100 repeat blocks were isolated. The inserts from the clones
CC were amplified by primers T65798-T66047. Those clones where the repeat
CC sequence has been determined are shown in T65704-797. This repeat
CC sequence is from the marker clone Md39 which contains the repeat
CC sequence having the formula: (TC)7(CA)23.
SQ Sequence 91 BP; 20 A; 43 C; 1 G; 27 T;

CC CA repeats separated by no more than 3 consecutive non-repeat bases;
CC 3 compound perfect repeats which are uninterrupted runs of CA separated
CC by no more than 3 consecutive non-repeat bases from a run of at least
CC 5 uninterrupted dinucleotide or longer repeats of a sequence other than
CC (dc-da)n.(dg-dt)n, or from at least 10 uninterrupted mononucleotides; and
CC 4) imperfect compound repeats which are defined as for the perfect
CC compound repeats except that the runs of CA are interrupted.
CC This sequence is an example of a compound imperfect repeat sequence of
CC structure T(CT)12GTT(TC)11(CA)14A(CA)5.
SQ Sequence 92 BP; 21 A; 43 C; 1 G; 27 T;

Query Match 0.2%; Score 50; DB 1; Length 92;
Best Local Similarity 100.0%; Pred. No. 2.9e-07;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3332 CTCCTCTCTCTCTCTCTCTCTCTCACACACACACACACACACACA 3981
Db 30 CTCCTCTCTCTCTCTCTCTCTCTCACACACACACACACACACACA 79

RESULT 4
T65707
ID T65707 standard; DNA; 56 BP.
AC T65707;
DT 17-JUN-1997 (first entry)
DE Repeat sequence from polymorphic marker clone Md5.
KM Polymorphism; repeat sequence; genetic marker; primer; amplification;
KM PCR; polymerase chain reaction; paternity; maternity; human; pedigree;
KM linkage analysis; genetic disease; animal; plant; breeding; locus;
KM hybridisation; chromosome; ds.
OS Homo sapiens.
PN US5582979-A.
PD 10-DEC-1996.
PF 21-APR-1989; 341562.
PR 21-APR-1989; US-341562.
PR 05-SEP-1991; US-754351.
PR 04-APR-1994; US-222177.
PA (MARS-) MARSHFIELD CLINIC.
PI Weber JL.
DR WPI; 97-042299/04.
PT Detection of polymorphic genetic markers of the form
PT (dc-da)n(dg-dt)n - using novel nucleic acid mols. as primers
PS Claim 1; Column 9-10; 186pp; English.
CC The invention relates to the isolation of polymorphic repeat sequences
CC having the sequence (dc-da)n.(dg-dt)n which can be used as genetic
CC markers. Primers based on these sequences can be used to detect these
CC repeats, especially for use in e.g. paternity or maternity testing,
CC human genetic analysis such as linkage analysis of genetic disease,
CC commercial animal or plant breeding or pedigree analysis. Clones
CC containing the repeat sequences were isolated by hybridisation of
CC chromosome-specific phage libraries with a synthetic poly(dc-da).(dg-dt)
CC probe. Over 100 repeat blocks were isolated. The inserts from the clones
CC were amplified by primers T65798-T66047. Those clones where the repeat
CC sequence has been determined are shown in T65704-797. This repeat
CC sequence is from the marker clone Md39 which contains the repeat
CC sequence having the formula: (CT)7(CA)23.
SQ Sequence 56 BP; 23 A; 28 C; 0 G; 5 T;

Query Match 0.1%; Score 42; DB 1; Length 56;
Best Local Similarity 100.0%; Pred. No. 0.00012;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3344 CTCCTCTCTCTCTCTCTCTCTCTCACACACACACACACACACACA 3985
Db 1 CTCCTCTCTCTCTCTCTCTCTCTCACACACACACACACACACACA 42

RESULT 5
T65791
ID T65791 standard; DNA; 50 BP.
AC T65791;

```

DE DE-Repeat sequence from polymorphic marker clone Mfd120.
KW polymorphism; repeat sequence; genetic marker; primer; amplification;
KW PCR; polymerase chain reaction; paternity; maternity; human; pedigree;
KW linkage analysis; genetic disease; animal; plant; breeding; locus;
KW hybridisation; chromosome; ds.
OS Homo sapiens.
PN US5582979-A.
PD 10-DEC-1996.
PF 21-APR-1989; 341562.
PR 05-SEP-1991; US-754351.
PR 04-APR-1994; US-222177.
PA (MARS-) MARSHFIELD CLINIC.
PI Weber JL;
DR WPI: 97-042299/04.
PT Detection of polymorphic genetic markers of the form
PS (dc-da)n(gc-dt)n - using novel nucleic acid mols. as primers
PS Claim 1; Column 13-14; 18pp; English.
CC The invention relates to the isolation of polymorphic repeat sequences
CC having the sequence (dc-da)n.(dg-dt)n which can be used as genetic
CC markers. Primers based on these sequences can be used to detect these
CC repeats, especially for use in e.g. paternity or maternity testing,
CC human genetic analysis such as linkage analysis of genetic disease,
CC commercial animal or plant breeding or pedigree analysis. Clones
CC containing the repeat sequences were isolated by hybridisation of
CC chromosome-specific phage libraries with a synthetic poly(dc-da).(dg-dt)
CC probe. Over 100 repeat blocks were isolated. The inserts from the clones
CC were amplified by primers T65798-T66047. Those clones where the repeat
CC sequence has been determined are shown in T65704-T6797. This repeat
CC sequence is from the marker clone Mdf120 which contains the repeat
SQ Sequence 50 BP; 20 A; 25 C; 0 G; 5 T;

Query Match 0.1%; Score 41; DB 1; Length 50;
Best Local Similarity 100.0%; Pred. No. 0.00025;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0

OY 3945 TCTCTCTCAGACACACACACACACACACACACA 3985
DB 1 TCTCTCTCAGACACACACACACACACACACACA 41
|||||
RESULT 6
ID Q33603 standard; DNA; 56 BP.
AC Q33603;
DT 02-FEB-1993 (first entry)
DE Microsatellite sequence from clone AGLA296.
KM PCR; selection; primers; OPRIPIM; breeding; cattle; parentage;
KM genetic mapping; traits; amplification; ss.
OS Bos taurus.
PN WO9213102-A.
PD 06-AUG-1992.
PF 15-JAN-1992; U00340.
PR 15-JAN-1991; US-642342.
PA (GENM-) GENMARK.
PI Georges M, Massey JM;
PT Polymorphic bovine DNA markers - used in genetic identification,
PT gene mapping, and selective breeding
PS Table 7; Page 170; 517pp; English.
CC The sequence is that of a bovine microsatellite sequence obtd. by
CC screening a library of bovine MbOI DNA fragments of between
CC 250 and 500 bp with an (AC)15 and a (TC)15 oligonucleotide probe.
CC One out of 50 clones cross-hybridised. Assuming independent
CC distribution of microsatellites and MbOI sites, the frequency of
CC (T6)n >9 microsatellites in the bovine genome is estimated at >100,
CC 000. The sequence information for ca. 230 such bovine microsatelites
CC is summarised in the specification and indexed herein (see below).
CC The sequences upstream and downstream of the microsatellite sequence
CC were used to generate the required PCR primers for in vitro

```

```

CC amplification of the corresp. microsatellite (using the program
CC OPTIPRIM). The microsatellites may be used to identify individuals,
CC for percentage testing, and in the genetic mapping of economic trait
CC loci, or genes involved the determinism of economically important
CC traits esp. in cattle, to allow selective breeding.
CC See also Q33501-34437.
SQ Sequence 56 BP; 19 A; 0 C; 28 G; 9 T;

Query Match 0.1%; Score 41; DB 1; Length 56;
Best Local Similarity 100.0%; Pred. No. 0.00024; Indels 0; Gaps 0;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33932 CTCCTCTCTCTCTCTCTCTCTCTCACACACACACACACAC 33972
|||||
DB 41 CTCCTCTCTCTCTCTCTCTCTCTCACACACACACACACACAC 1

RESULT 7
Q33779/c
ID Q33779 standard; DNA: 62 BP.
AC Q33779;
DT 02-FEB-1993 (first entry)
DE Microsatellite sequence from clone TGLA182.
KW PCR; selection; primers; OPTIPRIM; breeding; cattle; parentage;
KW genetic mapping; traits; amplification; ss.
OS Bos taurus.
FN WO9213102-A.
PD 06-AUG-1992.
PF 15-JAN-1992; U00340.
PR 15-JAN-1991; US-642342.
PI (GENEM-) GENMARK.
PA Georges M. Massey JM;
DR WPI; 92-284684/34.
PT Polymorphic bovine DNA markers - used in genetic identification,
PT gene mapping, and selective breeding
PT Table 7; Page 241; 517pp; English.
PS The sequence is that of a bovine microsatellite sequence obd. by
CC screening a library of bovine MboI DNA fragments of between
CC 250 and 500 bp with an (AC)15 and a (TC)15 oligonucleotide probe.
CC One out of 50 clones cross-hybridised. Assuming independent
CC distribution of microsatellites and MboI sites, the frequency of
CC (T6)n >9 microsatellites in the bovine genome is estimated at >100,
CC 000. The sequence information for ca. 230 such bovine microsatellites
CC is summarised in the specification and indexed herein (see below).
CC The sequences upstream and downstream of the microsatellite sequence
CC were used to generate the required PCR primers for in vitro
CC amplification of the corresp. microsatellite (using the program
CC OPTIPRIM). The microsatellites may be used to identify individuals,
CC for percentage testing, and in the genetic mapping of economic trait
CC loci, or genes involved the determinism of economically important
CC traits esp. in cattle, to allow selective breeding.
CC See also Q33501-34437.
SQ Sequence 62 BP; 22 A; 0 C; 31 G; 9 T;

Query Match 0.1%; Score 41; DB 1; Length 62;
Best Local Similarity 100.0%; Pred. No. 0.00024; Indels 0; Gaps 0;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33932 CTCCTCTCTCTCTCTCTCTCTCTCACACACACACACAC 33972
|||||
DB 41 CTCCTCTCTCTCTCTCTCTCTCTCACACACACACACACACAC 1

RESULT 8
Q33594/c
ID Q33594 standard; DNA: 63 BP.
AC Q33594;
DT 02-FEB-1993 (first entry)
DE Microsatellite sequence from clone AGLA29.
KW PCR; selection; primers; OPTIPRIM; breeding; cattle; parentage;
KW genetic mapping; traits; amplification; ss.

```

Bos taurus.

M09213102-A.
06-APR-1987
15-JAN-1992; U00340.
15-JAN-1991; US-642342.
(GENM-) GENMARK.
PI Georges M. Massey JM;
DI WPI: 92-284684/34.
PT Polymorphic bovine DNA markers - used in genetic identification,
Pr gene mapping, and selective breeding

PS Table 7, Page 167, 517pp: English.

The sequence is that of a bovine microsatellite sequence obt'd. by screening a library of bovine MbOI DNA fragments of between 250 and 500 bp with an (AC)15 and a (TC)15 oligonucleotide probe. One out of 50 clones cross-hybridised. Assuming independent distribution of microsatellites and MbOI sites, the frequency of (T6)n >g microsatellites in the bovine genome is estimated at >100,000. The sequence information for ca. 230 such bovine microsatellites is summarised in the specification and indexed herein (see below).

The sequences upstream and downstream of the microsatellite sequence were used to generate the required PCR primers for in vitro amplification of the corresp. microsatellite (using the program OPTPRIM). The microsatellites may be used to identify individuals, for parentage testing, and in the genetic mapping of economic trait loci, or genes involved in the determination of economically important traits esp. in cattle, to allow selective breeding.

See also Q33501-34437.

SQ Sequence 63 BP; 22 A; 0 C; 32 G; 9 T;

Query Match 0.1%; Score 41; DB 1; Length 63;
Best Local Similarity 100.0%; Pred.No. 0.00024;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3932 CTCCTCCTCTCTCTCTCTCTCACACACACACACACAC 3972
DB 41 CTCCTCCTCTCTCTCTCTCTCTCATACACACACACACAC 1

|||||
|||

RESULT 9
O33618/c
ID O33618 standard; DNA: 40 BP.
AC O33618;
DT 02-FEB-1993 (first entry)
DE Microsatellite sequence from clone AGIAB.
KW PCR selection; primers: OPTPRIM; breeding; cattle; parentage;
KM genetic mapping; trials; amplification; ss.
OS Bos taurus.
PN WO9213102-A.
PD 06-APR-1987.
PE 15-JAN-1992; U00340.
PR 15-JAN-1991; US-642342.
PA (GENM-) GENMARK.
PI Georges M. Massey JM;
DI WPI: 92-284684/34.
DR Polymorphic bovine DNA markers - used in genetic identification,
Pr gene mapping, and selective breeding

PS Table 7, Page 176, 517pp: English.

The sequence is that of a bovine microsatellite sequence obtained by screening a library of bovine MbOI DNA fragments of between 250 and 500 bp with an (AC)15 and a (TC)15 oligonucleotide probe. One out of 50 clones cross-hybridized. Assuming independent distribution of microsatellites and MbOI sites, the frequency of (T6)n >g microsatellites in the bovine genome is estimated at >100,000. The sequence information for ca. 230 such bovine microsatellites is summarized in the specification and indexed herein (see below).

The sequences upstream and downstream of the microsatellite sequence were used to generate the required PCR primers for in vitro amplification of the corresp. microsatellite (using the program OPTPRIM). The microsatellites may be used to identify individuals, for parentage testing, and in the genetic mapping of economic trait loci, or genes involved in the determination of economically important traits esp. in cattle, to allow selective breeding.

```
CC See also Q33501-34437.
```

```
SQ Sequence 40 BP; 11 A; 0 C; 20 G; 9 T;
```

```
Query Match          0.1%; Score 40; DB 1; Length 40;  
Best Local Similarity 100.0%; Pred. No. 0.00055;  
Matches    40; Conservative   0; Mismatches    0; Indels      0; Gaps      0;
```

```
OY 3932 CTCCTCCTCTCTCTCTCTCTCACACACACACACACA 3971  
         |||||||  
DB     40 CTCTCTCTCTCTCTCTCTCTCAACACACACACACACA 1
```

```
RESULT 10  
ID V19044/c       VI9044 standard; DNA: 40 BP.  
AC V19044:  
DT 28-JUL-1998 (first entry)  
DE Alu PCR primer 1.  
PR PCR; primer; amplification; Alu repeat sequence; vector;  
KW circular yeast artificial chromosome; YAC; ss.  
OS Synthetic.  
OI Saccharomyces sp.  
PN MO9601573-A1.  
PD 15-JAN-1998.  
PE 09-JUL-1998; U11478.  
PF 09-JUL-1996; WO-011478.  
PI (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
DR Kourina NY, Lartionov VL, Perkins EL, Resnick MA;  
WPI: 98-110234/10.  
PT Preparation of yeast artifical chromosomes - by in vivo  
recombination using vector comprising yeast centromere, marker,  
yeast telomere and nucleic acid for recombination  
PS Example 1; Page 45; 11/Pp; English.  
CC This is the nucleotide sequence for the PCR primer used in the  
amplification of the Alu repeat sequence, which is used to  
demonstrate the processes described in the invention. It involves  
the creation and use of circular yeast artifical chromosome (YAC)  
to selectively clone specific nucleic acids from a background of  
mixed nucleic acids by introducing the vector(s) into E. coli cells.  
CC They can be used to rapidly isolate human DNA where only a part of the  
sequence of DNA is known. Using the methods large fragments of DNA can  
be easily cloned and analysed.  
SQ Sequence 40 BP; 7 A; 12 C; 13 G; 8 T;
```

```
Query Match          0.1%; Score 40; DB 1; Length 40;  
Best Local Similarity 100.0%; Pred. No. 0.00055;  
Matches    40; Conservative   0; Mismatches    0; Indels      0; Gaps      0;
```

```
OY 24755 TCGGGCTCCCAGTAGTGCGATTACAGGCCGTGCACC 24794  
        |||||  
DB     40 TCGGCTCCCCAAGTGTGGATTTCAGGCGTTGCCACC 1
```

```
RESULT 11  
ID O95218/c       O95218 standard; DNA: 85 BP.  
AC O95218:  
DT 08-FEB-1996 (first entry)  
DE Simple tandem repeat (STR) corresponding to the 2nd part of wg3bl0.  
PR Simple tandem repeat; STR; wg3bl0; treatment; genetic; diagnosis;  
KW characterisation; mapping; linkage studies; analysts; alleles;  
KM second part; ss.  
OS Synthetic.  
PN WO9517522-A2.  
PD 29-JUN-1995.  
PE 21-DEC-1994; GO2789.  
PF 21-DEC-1993; GB-026052.  
PI (UYLE-) UNITV LEICESTER.  
DR Armour J, Jeffreys AJ;  
WPI: 95-240682/31.  
PT Identifying simple tandem repeat loci in DNA - by screening DNA
```


CC synthetic polyester genes given in 071801-02. These fragments were
CC used in the production of the full length polyester gene in the plasmid
CC pMAL-p2. This plasmid was used so that the polyester could be expressed
CC in an E. coli cell free medium, as pMAL-p2 is recognized by E. coli
CC ribozymes and contains the necessary initiator sequences and the
CC template for the polypeptide leader. The resulting plasmid pMAL-2c
CC contains the complete polyester coding sequence ligated in frame with
CC the leader sequence. The expressed sequences represent synthetic
CC genes which were used in the synthesis of polyesters derived from the
CC a-amino acid analogues of an amino acid except Pro. Of the three stop
CC codons, UAA, UAG and UGA, one of these stop codons is reversed for the
CC stop signal for the polyester. The remaining two stop codons are
CC available for encoding the monomer units of the polyester. Lactate
CC and glycolate are encoded by UAA and UAG, respectively. Initiation of
CC translation always starts at an AUG codon, therefore the first lactate
CC of the chain may be encoded by an AUG codon. This is accomplished
CC through the use of a specially modified Met-tRNAf modified to carry
CC lactate instead of Met. Alternatively, a Met is incorporated in the
CC first position of the polymer chain. The resulting polymer is treated
CC with cyanogen bromide to remove the Met. Co-polymers produced in this
CC manner may be used to produce sutures, staples, clips, drug delivery
CC devices, pins and screws.
SQ Sequence 81 BP; 47 A; 6 C; 4 G; 24 T;

Query Match 0.1%; Score 39; DB 1; Length 81;
Best Local Similarity 100.0%; Pred. No. 0.00098;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10446 AAT 10484
|||||
DB 2 AAT 40

RESULT 15
071809
ID Q71809 standard; DNA; 103 BP.
AC Q71809;
DE 30-MAR-1995 (first entry)
DE PSORT-IC synthetic polyester gene 5' fragment and trpA initiator.
KW Synthetic gene; synthesis; polyester; a-amino acid analogue; clips;
KW stop codon; stop signal; monomer unit; lactate; glycolate; staples;
KW translation initiation; modified; Met-tRNAf; cyanogen bromide; pins;
KW co-polymer; sutures; drug delivery device; screws; ds.
OS Synthetic.
FH Key
FT misc-signal location/Qualifiers
FT 1..18
FT /*tag= a
FT /*note= "trpA initiator"
FT misc-feature 19..21
FT /*tag= b
FT /*note= "Start codon"
FT 22..93
FT /*tag= c
FT /*note= "Polyester gene 5' fragment"

CA2112716-A.
PD 01-JUL-1994.
PF 31-DEC-1993; 112716.
PR 31-DEC-1993; US-999520.
PA (USCU) US SURGICAL CORP.
PI Gruskin EA.
DR WPI, 94-272324/34.
PT New biocompatible medical devices and articles - using a lactic
PT acid-glycolic acid polyester in which monomer units are
PT individually and specifically determined
PS Example 1; Page 45; 67pp; English.
CC This sequence represents a fragment of the plasmid pSPORT IC which
CC contains the 5' fragment of the synthetic genes which encode polyesters,
CC given in 071801-02, under the control of the E. coli trpA transcription
CC initiator and followed by PstI and HindIII restriction sites. A further
CC fragment of the synthetic genes were ligated into pSPORT IC to give
CC plasmid pSPORT id (see also 071810-11), with the final fragment being
CC inserted in a third round of ligation to give pSPORT Ie (see also 071812-

CC 13). The E. coli trpA initiator sequence was used to express the full
CC length sequences given in 071801-02 in an E. coli cell free medium. The
CC expressed sequences represent synthetic genes which were used in the
CC synthesis of polyesters derived from the a-amino acid analogues of an
CC amino acid except Pro. Of the three stop codons, UAA, UAG and UGA, one
CC of these stop codons is reversed for the stop signal for the polyester.
CC The remaining two stop codons are available for encoding the monomer
CC units of the polyester. Lactate and glycolate are encoded by UAA and
CC UAG, respectively. Initiation of translation always starts at an AUG
CC codon, therefore the first lactate of the chain may be encoded by an
CC AUG codon. This is accomplished through the use of a specially modified
CC Met-tRNAf modified to carry lactate instead of Met. Alternatively, a
CC Met is incorporated in the first position of the polymer chain. The
CC resulting polymer is treated with cyanogen bromide to remove the Met.
CC Co-polymers produced in this manner may be used to produce sutures,
CC staples, clips, drug delivery devices, pins and screws.
SQ Sequence 103 BP; 58 A; 7 C; 29 T;

Query Match 0.1%; Score 39; DB 1; Length 103;
Best Local Similarity 100.0%; Pred. No. 0.00093;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10446 AAT 10484
|||||
DB 23 AAT 61

Search completed: April 1, 2000, 04:18:02
Job time: 260880 sec

Mon Apr 3 08:24:13 2000

us-08-852-495c-1_copy_190000_220000.png

Page 7

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 31, 2000, 19:15:35 ; Search time 339.54 Seconds
(Without alignments)
10578.471 Million cell updates/sec

Title: US-08-852-495c-1_COPY_190000_220000
Perfect score: 30001
Sequence: 1 GATATATCATGTCGTGACTG.....AGGGCAAGAGATCATG 30001

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 214294 seqs, 59861574 residues

Word size : 0

Total number of hits satisfying chosen parameters: 349160

Minimum DB seq length: 8

Maximum DB seq length: 105

Post-processing: Listing first 45 summaries

Database : Issued_Patents_NA:*

- 1: /cgn2_6/ptodata/1/lna/5A_COMB.seq:*
- 2: /cgn2_6/ptodata/1/lna/5B_COMB.seq:*
- 3: /cgn2_6/ptodata/1/lna/5C_COMB.seq:*
- 4: /cgn2_6/ptodata/1/lna/5D_COMB.seq:*
- 5: /cgn2_6/ptodata/1/lna/6_COMB.seq:*
- 6: /cgn2_6/ptodata/1/lna/PCTU59_COMB.seq:*
- 7: /cgn2_6/ptodata/1/lna/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	54	0.2	60	1	US-08-222-177A-244 Sequence 244, App
2	50	0.2	91	1	US-08-222-177A-166 Sequence 166, App
3	50	0.2	92	1	US-08-222-177A-430 Sequence 430, App
4	42	0.1	56	1	US-08-222-177A-65 Sequence 65, App
5	41	0.1	50	1	US-08-222-177A-379 Sequence 379, App
6	40	0.1	85	3	US-08-332-766A-44 Sequence 44, App
7	37	0.1	60	3	US-08-454-557C-57 Sequence 57, App
8	37	0.1	60	4	US-08-340-426D-57 Sequence 57, App
9	37	0.1	60	6	US-08-450-673C-57 Sequence 57, App
10	37	0.1	60	6	PCT-US95-17111A-57 Sequence 57, App
11	37	0.1	76	3	US-08-454-557C-69 Sequence 69, App
12	37	0.1	76	4	US-08-340-426D-69 Sequence 69, App
13	37	0.1	76	4	US-08-450-673C-69 Sequence 69, App
14	37	0.1	76	6	PCT-US95-17111A-69 Sequence 69, App
15	35	0.1	85	3	US-08-454-557C-92 Sequence 92, App
16	35	0.1	85	4	US-08-340-426D-92 Sequence 92, App
17	35	0.1	85	4	US-08-450-673C-92 Sequence 92, App
18	35	0.1	85	6	PCT-US95-17111A-92 Sequence 92, App
19	33	0.1	33	1	US-08-222-177A-312 Sequence 312, App
20	33	0.1	40	1	US-08-222-177A-175 Sequence 175, App
21	33	0.1	46	1	US-08-222-177A-71 Sequence 71, App
22	33	0.1	51	1	US-08-222-177A-317 Sequence 317, App
23	33	0.1	54	2	US-08-469-802B-27 Sequence 27, App
24	33	0.1	54	3	US-08-267-803B-45 Sequence 45, App
25	33	0.1	62	7	5506118-3 Patent No. 5506118
26	32	0.1	33	1	US-08-222-177A-59 Sequence 59, App

27	32	0.1	33	1	US-08-222-177A-218 Sequence 218, App
28	32	0.1	34	1	US-08-222-177A-110 Sequence 110, App
29	32	0.1	34	1	US-08-222-177A-172 Sequence 172, App
30	32	0.1	34	1	US-08-222-177A-180 Sequence 180, App
31	32	0.1	34	1	US-08-222-177A-192 Sequence 192, App
32	32	0.1	34	1	US-08-222-177A-322 Sequence 322, App
33	32	0.1	35	1	US-08-222-177A-77 Sequence 77, App
34	32	0.1	35	1	US-08-222-177A-189 Sequence 189, App
35	32	0.1	36	1	US-08-222-177A-104 Sequence 104, App
36	32	0.1	36	1	US-08-222-177A-358 Sequence 358, App
37	32	0.1	37	1	US-08-222-177A-140 Sequence 140, App
38	32	0.1	38	1	US-08-222-177A-198 Sequence 198, App
39	32	0.1	38	1	US-08-222-177A-397 Sequence 397, App
40	32	0.1	39	1	US-08-222-177A-137 Sequence 137, App
41	32	0.1	40	1	US-08-222-177A-119 Sequence 119, App
42	32	0.1	40	1	US-08-222-177A-400 Sequence 400, App
43	32	0.1	40	1	US-08-222-177A-403 Sequence 403, App
44	32	0.1	41	1	US-08-222-177A-74 Sequence 74, App
45	32	0.1	41	1	US-08-222-177A-183 Sequence 183, App

ALIGNMENTS

RESULT 1
US-08-222-177A-244
; Sequence 244, Application US/08222177A
; Patent No. 5582979
; GENERAL INFORMATION:
; APPLICANT: Weber, James L.
; TITLE OF INVENTION: LENGTH POLYMORPHISMS IN
; TITLE OF INVENTION: (dc-da)n.(dg-dt)n SEQUENCES AND METHODS OF USING SAME
; NUMBER OF SEQUENCES: 460
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dewalt Ross & Stevens, S.C.
; STREET: 8000 Excelstor Drive, Suite 401
; City: Madison
; STATE: Wisconsin
; COUNTRY: USA
; ZIP: 53717-1914
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/222,177A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/341,562
; FILING DATE: 21-APR-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Sara, Charles S.
; REGISTRATION NUMBER: 30,492
; REFERENCE/DOCKET NUMBER: 09865.601
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (608) 831-2100
; TELEFAX: (608) 831-2106
; TELEX:
; INFORMATION FOR SEQ ID NO: 244:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 60 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; IMMEDIATE SOURCE:
; CLONE: mfd67rs
; US-08-222-177A-244
Query Match 0.2%; Score 54; DB 1; Length 60;

Patent No. 5830670
GENERAL INFORMATION:
APPLICANT: de la Monte, Suzanne
TITLE OF INVENTION: Neutral Thread Protein Gene Expression and Detection
TITLE OF INVENTION: of Alzheimer's Disease
NUMBER OF SEQUENCES: 121
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/454,557C
FILING DATE: 30-MAY-1995
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609.3840003
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 57:
SEQUENCE CHARACTERISTICS:
LENGTH: 60 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
US-08-454-557C-57

Query Match 0.1%; Score 37; DB 3; Length 60;
Best Local Similarity 100.0%; Pred. No. 0.0018;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21188 GCTAATTTTGTATTGTAGTAGACAGAGGTTTCAC 21224
DB 11 GCTAATTTTGTATTGTAGTAGACAGAGGTTTCAC 47

RESULT 8
US-08-340-426D-57
Sequence 57, Application US/08340426D
Patent No. 5948634
GENERAL INFORMATION:
APPLICANT: de la Monte, Suzanne
APPLICANT: Wands, Jack R.
TITLE OF INVENTION: Neutral Thread Protein Gene Expression and Detection
TITLE OF INVENTION: of Alzheimer's Disease
NUMBER OF SEQUENCES: 121
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/340,426D
FILING DATE: 14-NOV-1994

CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609.3840002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 57:
SEQUENCE CHARACTERISTICS:
LENGTH: 60 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
US-08-340-426D-57

Query Match 0.1%; Score 37; DB 4; Length 60;
Best Local Similarity 100.0%; Pred. No. 0.0018;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21188 GCTAATTTTGTATTGTAGTAGACAGAGGTTTCAC 21224
DB 11 GCTAATTTTGTATTGTAGTAGACAGAGGTTTCAC 47

RESULT 9
US-08-450-673C-57
Sequence 57, Application US/08450673C
Patent No. 5948868
GENERAL INFORMATION:
APPLICANT: de la Monte, Suzanne
APPLICANT: Wands, Jack R.
TITLE OF INVENTION: Neutral Thread Protein Gene Expression and Detection
TITLE OF INVENTION: of Alzheimer's Disease
NUMBER OF SEQUENCES: 121
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/450,673C
FILING DATE: 30-MAY-1995
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609.3840004
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 57:
SEQUENCE CHARACTERISTICS:
LENGTH: 60 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
US-08-450-673C-57

Query Match 0.1%; Score 37; DB 4; Length 60;
Best Local Similarity 100.0%; Pred. No. 0.0018;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21188 GCTAATTTTGTATTGTAGTAGACAGAGGTTTCAC 21224
DB 11 GCTAATTTTGTATTGTAGTAGACAGAGGTTTCAC 47

DB 11 GCTAATTTTGTATTAGTAGACAGCGTTTCAC 47

RESULT 10
PCT-US95-17111A-57
Sequence 57, Application PC/TUS9517111A
GENERAL INFORMATION:
APPLICANT: de la Monte, Suzanne
TITLE OF INVENTION: Neural Thread Protein Gene Expression and
TITLE OF INVENTION: Detection of Alzheimer's Disease
NUMBER OF SEQUENCES: 121
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/17111A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/340,426
FILING DATE: 14-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609,3840002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 57:
SEQUENCE CHARACTERISTICS:
LENGTH: 60 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
PCT-US95-17111A-57

Query Match 0.1%; Score 37; DB 6; Length 60;
Best Local Similarity 100.0%; Pred. No. 0.0018;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21188 GCTAATTTTGTATTAGTAGACAGCGTTTCAC 21224
DB 11 GCTAATTTTGTATTAGTAGACAGCGTTTCAC 47

RESULT 11
US-08-454-557C-69
Sequence 69, Application US/08454557C
Patent No. 5830670
GENERAL INFORMATION:
APPLICANT: de la Monte, Suzanne
TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection
TITLE OF INVENTION: of Alzheimer's Disease
NUMBER OF SEQUENCES: 121
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.

ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/454,557C
FILING DATE: 30-MAY-1995
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609,3840003
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 69:
SEQUENCE CHARACTERISTICS:
LENGTH: 76 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
US-08-454-557C-69

Query Match 0.1%; Score 37; DB 3; Length 76;
Best Local Similarity 100.0%; Pred. No. 0.0017;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21188 GCTAATTTTGTATTAGTAGACAGCGTTTCAC 21224
DB 11 GCTAATTTTGTATTAGTAGACAGCGTTTCAC 47

RESULT 12
US-08-340-426D-69
Sequence 69, Application US/08340426D
Patent No. 5948634
GENERAL INFORMATION:
APPLICANT: de la Monte, Suzanne
TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection
TITLE OF INVENTION: of Alzheimer's Disease
NUMBER OF SEQUENCES: 121
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/340,426D
FILING DATE: 14-NOV-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609,3840002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 69:
SEQUENCE CHARACTERISTICS:
LENGTH: 76 base pairs
TYPE: nucleic acid
STRANDEDNESS: both

TOPOLOGY: both
US-08-340-426D-69

Query Match 0.1%; Score 37; DB 4; Length 76;
Best Local Similarity 100.0%; Pred. No. 0.0017;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 21188 GCTAATTTTGTATTTTGTAGACAGAGGTTTCAC 21224
|||||
DB 11 GCTAATTTTGTATTTTGTAGACAGAGGTTTCAC 47

RESULT 13
US-08-450-673C-69
Sequence 69, Application US/08450673C
Patent No. 5948888

GENERAL INFORMATION:

APPLICANT: de la Monte, Suzanne

APPLICANT: Wands, Jack R.

TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection

TITLE OF INVENTION: of Alzheimer's Disease

NUMBER OF SEQUENCES: 121

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.

STREET: 1100 New York Avenue, Suite 600

CITY: Washington

STATE: D.C.

COUNTRY: U.S.A.

ZIP: 20005-3934

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/450,673C

FILING DATE: 30-MAY-1995

CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:

NAME: Ludwig, Steven R.

REGISTRATION NUMBER: 36,203

REFERENCE/DOCKET NUMBER: 0609.3840004

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 371-2600

TELEFAX: (202) 371-2540

INFORMATION FOR SEQ ID NO: 69:

SEQUENCE CHARACTERISTICS:

LENGTH: 76 base pairs

TYPE: nucleic acid

STRANDEDNESS: both

TOPOLOGY: both

US-08-450-673C-69

Query Match 0.1%; Score 37; DB 4; Length 76;
Best Local Similarity 100.0%; Pred. No. 0.0017;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 21188 GCTAATTTTGTATTTTGTAGACAGAGGTTTCAC 21224
|||||
DB 11 GCTAATTTTGTATTTTGTAGACAGAGGTTTCAC 47

RESULT 14
PCT-US95-17111A-69
Sequence 69, Application PC/TUS9517111A

GENERAL INFORMATION:

APPLICANT: de la Monte, Suzanne

APPLICANT: Wands, Jack R.

TITLE OF INVENTION: Neural Thread Protein Gene Expression and

TITLE OF INVENTION: Detection of Alzheimer's Disease

NUMBER OF SEQUENCES: 121

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.

STREET: 1100 New York Avenue, Suite 600

CITY: Washington

STATE: D.C.

COUNTRY: U.S.A.

ZIP: 20005-3934

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US95/17111A

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/340,426

FILING DATE: 14-NOV-1994

ATTORNEY/AGENT INFORMATION:

NAME: Ludwig, Steven R.

REGISTRATION NUMBER: 36,203

REFERENCE/DOCKET NUMBER: 0609.3840002

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 371-2600

TELEFAX: (202) 371-2540

INFORMATION FOR SEQ ID NO: 69:

SEQUENCE CHARACTERISTICS:

LENGTH: 76 base pairs

TYPE: nucleic acid

STRANDEDNESS: both

TOPOLOGY: both

PCT-US95-17111A-69

OY 21188 GCTAATTTTGTATTTTGTAGACAGAGGTTTCAC 21224
|||||
DB 11 GCTAATTTTGTATTTTGTAGACAGAGGTTTCAC 47

RESULT 15
US-08-454-557C-92
Sequence 92, Application US/08454557C
Patent No. 5850670

GENERAL INFORMATION:

APPLICANT: de la Monte, Suzanne

APPLICANT: Wands, Jack R.

TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection

TITLE OF INVENTION: of Alzheimer's Disease

NUMBER OF SEQUENCES: 121

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.

STREET: 1100 New York Avenue, Suite 600

CITY: Washington

STATE: D.C.

COUNTRY: U.S.A.

ZIP: 20005-3934

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/454,557C

FILING DATE: 30-MAY-1995

CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:

NAME: Ludwig, Steven R.

REGISTRATION NUMBER: 36,203


```

REFERENCE/DOCKET NUMBER: 0609.3840003
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 92:
SEQUENCE CHARACTERISTICS:
LENGTH: 85 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
US-08-454-557C-92

```

```

Query Match          0.1%; Score 35; DB 3; Length 85;
Best Local Similarity 100.0%; Pred. No. 0.0075;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 18907 GCCAGATGGTCTCGATCTCTGACCTGACCTGATCC 18941
      |||||||||||||||||||||||||||||||||||
Db 9 GCCAGATGGTCTCGATCTCTGACCTGACCTGATCC 43

```

Search completed: April 1, 2000, 04:07:52
 Job time: 260334 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 31, 2000, 13:42:06 ; Search time 8179.17 Seconds
(without alignments)
13849.045 Million cell updates/sec

Title: US-08-852-495c-1_COPY_190000_220000

Perfect score: 30001
Sequence: 1 GTATATATCATGCTACTG.....AGGGCAAAAGATCATG 30001

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 4538634 seqs, 1887831982 residues

Word size : 0

Total number of hits satisfying chosen parameters: 130662

Minimum DB seq length: 8

Maximum DB seq length: 105

Post-processing: Listing first 45 summaries

Database :
EST: *
1: em_est1: *
2: em_est2: *
3: em_est3: *
4: em_est4: *
5: em_est5: *
6: em_est6: *
7: em_est7: *
8: em_est8: *
9: em_est9: *
10: em_est10: *
11: em_est11: *
12: em_est12: *
13: em_est13: *
14: em_est14: *
15: em_est15: *
16: em_est16: *
17: em_est17: *
18: em_est18: *
19: em_est19: *
20: gb_est1: *
21: gb_est2: *
22: gb_est3: *
23: gb_est4: *
24: gb_est5: *
25: gb_est6: *
26: gb_est7: *
27: gb_est8: *
28: gb_est9: *
29: gb_est10: *
30: gb_est11: *
31: gb_est12: *
32: gb_est13: *
33: gb_est14: *
34: gb_est15: *
35: gb_est16: *
36: gb_est17: *
37: gb_est18: *
38: gb_est19: *
39: gb_est20: *
40: gb_est21: *
41: gb_est22: *
42: gb_est23: *
43: gb_est24: *

44: gb_est25: *
45: gb_est26: *
46: gb_est27: *
47: gb_est28: *
48: gb_est29: *
49: gb_est30: *
50: gb_est31: *
51: gb_est32: *
52: gb_est20: *
53: em_est21: *
54: em_est22: *
55: em_est23: *
56: em_est24: *
57: em_est25: *
58: em_est26: *
59: gb_est33: *
60: gb_est34: *
61: gb_est35: *
62: gb_est36: *
63: gb_est37: *
64: gb_est38: *
65: em_est27: *
66: em_est28: *
67: em_est29: *
68: em_est30: *
69: gb_est39: *
70: gb_est40: *
71: gb_est41: *
72: gb_est42: *
73: gb_est43: *
74: gb_est44: *
75: em_est31: *
76: em_est32: *
77: em_est33: *
78: em_est34: *
79: gb_gss1: *
80: gb_gss2: *
81: gb_gss3: *
82: gb_gss4: *
83: em_gss1: *
84: em_gss2: *
85: em_gss3: *
86: em_gss4: *
87: gb_gss5: *
88: gb_gss6: *
89: gb_gss7: *
90: gb_gss8: *
91: gb_gss9: *
92: em_gss5: *
93: em_gss6: *
94: em_gss7: *
95: em_gss8: *
96: em_gss9: *
97: em_gss10: *
98: em_gss11: *
99: gb_gss10: *
100: gb_gss11: *
101: em_gss12: *
102: gb_gss12: *
103: gb_gss13: *
104: gb_gss14: *
105: gb_gss15: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
c 1	48	0.2	102	81	B62983

B62983 CTR-HSP-669

```

c 2 48 0.2 103 81 B48914
c 3 47 0.2 91 50 F29274
c 4 45 0.1 74 33 AA415364
c 5 44 0.1 82 44 AI305512
c 6 44 0.1 91 27 AA053038
c 7 43 0.1 101 39 AA835205
c 8 41 0.1 90 61 AI864984
c 9 41 0.1 97 39 AA835411
c 10 41 0.1 104 29 AA129957
c 11 40 0.1 95 37 AA669571
c 12 40 0.1 96 22 R67088
c 13 39 0.1 73 51 AU072757
c 14 39 0.1 80 44 AI262095
c 15 39 0.1 93 34 AU038446
c 16 39 0.1 103 104 A0582186
c 17 39 0.1 105 81 B74378
c 18 38 0.1 38 42 AI088003
c 19 38 0.1 58 35 C25772
c 20 38 0.1 104 36 AA614379
c 21 37 0.1 95 34 AA457423
c 22 36 0.1 52 20 D25845
c 23 36 0.1 71 47 AI475290
c 24 36 0.1 95 25 N88930
c 25 36 0.1 101 100 AQ280224
c 26 36 0.1 102 81 B80126
c 27 36 0.1 102 100 AQ319270
c 28 36 0.1 105 63 AI973231
c 29 36 0.1 105 100 AQ276193
c 30 35 0.1 37 22 R70733
c 31 35 0.1 54 36 AA601314
c 32 35 0.1 59 28 AA082835
c 33 35 0.1 88 38 AA809831
c 34 35 0.1 92 35 AA548135
c 35 35 0.1 93 21 R09732
c 36 35 0.1 94 23 R93021
c 37 35 0.1 95 24 H61099
c 38 35 0.1 95 35 AA578401
c 39 35 0.1 101 34 AA494273
c 40 35 0.1 103 30 AA228795
c 41 35 0.1 103 38 AA807640
c 42 34 0.1 34 25 N77004
c 43 34 0.1 61 39 AA873656
c 44 34 0.1 74 45 AI335049
c 45 34 0.1 75 79 HSMC09A06

```

ALIGNMENTS

```

RESULT 1
LOCUS B62983 102 bp DNA GSS 21-JUN-1998
DEFINITION CIT-HSP-669P15.TP CIT-HSP Homo sapiens genomic clone 669P15,
B62983
ACCESSION B62983.1 GI:2636892
VERSION B62983.1
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 102)
AUTHORS Kim, U.-J., Adams, M.D. and Simon, M.I.
TITLE Determination of clone end sequences of human Bacterial Artificial
Chromosomes
JOURNAL Unpublished (1997)
COMMENT Contact: Ung-Jin Kim
Caltech Genome Research Lab
California Institute of Technology
Division of Biology, MS 147-75, Pasadena, CA 91125, USA
Tel: 626 796 7066
Fax: 626 395 4901
Email: ung@ash.tree.caltech.edu

```

```

FEATURES
source
1..102
/organism="Homo sapiens"
/db_xref="GDB:5491647"
/db_xref="taxon:9606"
/clone_id="CIT-HSP"
/sex="Male"
/cell_type="Sperm"
/note="Vector: pBelobAC11; site_1: HindIII; site_2:
HindIII"

```

```

BASE COUNT 30 a 30 c 21 g 21 t
ORIGIN

```

```

Query Match 0.28; Score 48; DB 81; Length 102;
Best Local Similarity 100.0%; Pred. No. 2e-05;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 17313 TTTTGTATTTTGTAGATGGGTTTCACATGTTGCCAGGATG 17360
|||||
DB 59 TTTTGTATTTTGTAGATGGGTTTCACATGTTGCCAGGATG 12
|||||

```

```

RESULT 2
LOCUS B48914 103 bp DNA GSS 08-APR-1999
DEFINITION RPCI11-4A12.TP RPCI-11 Homo sapiens genomic clone RPCI-11-4A12,
B48914
ACCESSION B48914.1 GI:2601151
VERSION B48914.1
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 103)
AUTHORS Adams, M.D., Rounsley, S.D., Field, C.E., Baas, S., Linher, K.,
Golden, K., Berry, K., Granger, D., Suh, E., Wible, C., de Jong, P. and
Venter, J.C.
TITLE Use of BAC End Sequences for Sequence-Ready Map Building
JOURNAL Unpublished (1997)
COMMENT Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: madams@tigr.org
Clones are derived from the human BAC library RPCI-11. For BAC
library availability, please contact Pieter de Jong
(pieter@dejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
Research Genetics (info@resgen.com). BAC end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
Seq primer: SP6
Class: BAC ends.

```

```

FEATURES
source
1..103
/organism="Homo sapiens"
/db_xref="GDB:7501163"
/db_xref="taxon:9606"
/clone="RPCI-11-4A12"
/clone_id="RPCI-11"
/sex="Male"
/cell_type="Lymphocytes"
/note="Vector: pBAC63.6; Site_1: EcoRI; Site_2: EcoRI;
RPCI11 Human Male BAC Library"

```


AUTHORS	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap .
TITLE	National Cancer Institute, Cancer Genome Anatomy Project (CGAP)
JOURNAL	Tumor Gene Index
COMMENT	Unpublished (1997)
	Contact: Robert Strausberg, Ph.D.
	ret. (201) 405-1550

[5',
TGTACGACATCTGACATGGGAGCGCGCCCTTTTTTTTTTTTTTTTTTTT
3']; double-stranded cDNA was ligated to Eco RI adaptors
[AATTCGATCCTTGG], digested with Not I and cloned into the
Not I and Eco RI sites of the modified pT73 vector.
Library constructed by Bob Barstead. "
14 a 36 c 27 g 24 t

Query Match	0.18;	Score 43;	DB 39;	Length 101;
Best Local Similarity	100.0%;	Pred. No. 0.00085;		
Matches 43;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

```
QY 24577 ATCTCGGCTCACTGCAGCTCCGGCTCCCGGGTTCACGCCATT 24613
      |||||
Db 48 ATCTCGGCTCACTGCAGCTCCGGCTCCCGGGTTCACGCCATT 90
```

RESULT	8		
LOCUS	AI864984		
DEFINITION	AI864984	90 bp	EST
ACCESSION	U00611.x1	NCI-GAP-Lym12	Homo sapiens cDNA clone IMAGE:241115
VERSION	AI864984		similar to contains Alu repetitive element; mRNA sequence.
KEYWORDS	AI864984.1	GI:5529091	EST.
SOURCE	human.		
ORGANISM	Homo sapiens		

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1 (bases 1 to 90)	NCI-CCAP	http://www.ncbi.nlm.nih.gov/ncicgap . National Cancer Institute, Cancer Genome Anatomy Project (CGAP). Tumor Gene Index Unpublished (1997)		On May 18, 1998 this sequence version replaced gi:3137794.

JOURNAL COMMENT

Unpublished (1997) On May 18, 1998 this sequence version replaced g1:3137794.

Contact: Robert Strausberg, Ph.D.

Tel.: (301) 496-1550

Email: Robert_Strausberg@nih.gov

Life Technologies catalog #: 11547-015

DNA Sequencing by: Washington University Genome Sequencing Center

clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/dbfp/image/image.html

Seq primer: -40UP from G1bco.	Location/Qualifiers	source
	1. .90	

```

/lab_host="DH10B"
/notes="Organ: lymph node; Vector: PCMV-SPOrE; Site: 1;
Sali: Site 2; NotI: Cloned unidirectionally; primer:
Oligo dT. Average insert size 1.25 kb. Life Technologies
catalog #: 11547-015"
BASE COUNT
ORIGIN
13 a          27 g          21 t
29 c

```

```

Query Match 0.1%; Score 41; DB 61; Length 90;
Best Local Similarity 100.0%; Pred. No. 0 0041;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0.
Qy 24577 ATCTGGGCTCACTGCAAGCTCCGCGCTCCGGGTTACAGCCA 24617
|||||
Db 50 ATCTGGGCTCACTGCAAGCTCCGCGCTCCGGGTTACAGCCA 90
|||||

```

RESULT	9			
AA835411				
LOCUS	AA835411	97 bp	mRNA	EST
DEFINITION	a772806.s1 Barstead spleen HPIRB2 Homo sapiens cDNA clone			
IMAGE:	113466 3' similar to contains Alu repetitive element; , mRNA			
ACCESSION	AA835411			
VERSION	AA835411			
KEYWORDS	EST.	GT:2909139		
SOURCE	human.			

REFERENCE
AUTHORS
1 (bases 1 to 97)
Hiller, L., Allen, M., Bowles, L., Dubuque, T., Gaisel, G., Jost, R., Krizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marr, M., Martin, J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F., Tharling, B., White, Y., Wylie, T., Waterston, R. and Wilson, R.
TITLE
JOURNAL
WashU-NCI human EST Project
Unpublished (1997)
COMMENT
On May 8, 1995 this sequence version replaced g1:801230.

Email: est@watson.wustl.edu
This clone is available royalty-free through LNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -40m13 fwd. ET from Amersham.

```

FEATURES
    source
        location/Qualifiers
            1..97
                /organism="Homo sapiens"
                /db_xref="taxon:9605"
                /clone="IMAGE:1413466"
                /clone_id="Barstead spleen HPLRB2"
                /sex="male"
                /dev_stage="adult, 17 years"
                /lab_host="DH10B"
                /note="Organ: spleen; Vector: pT73D-Pac (Pharmacia) with
a modified polylinker; Site_1: EcoRI; Site_2: NotI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer (5
TGTTCAGCATCTGAGTGAGGAGCGCGCCCTTTTTTTTTTTTTTTTTT
3'); double-stranded cDNA was ligated to Eco RI adaptors
[AAATCGATCCTTG], digested with Not I and cloned into the
Not I and Eco RI sites of the modified pT73 vector.
Library constructed by Bob Barstead."
BASE COUNT
    16 a
    28 c
    24 g
    29 t

```

```
Query Match      0.1%; Score 41; DB 39; Length 97;  
Best Local Similarity    100.0%; Pred. No. 0.0039;  
Matches   41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

RESULT	10		
LOCUS	AA129957		
DEFINITION	AA129957 104 bp mRNA	EST	27-NOV-1996
	z866h04.1 Stragelene lung carcinoma 937218 Homo sapiens cDNA clone		
	IMAGE:565111 5' similar to contains Alu repetitive element;; mRNA		
	sequence.		
ACCESSION	AA129957		
VERSION	AA129957.1	GI:1689679	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;		
	Eutheria; Primates; Catarrhini; Homnidae; Homo.		

REFERENCE 1 (bases 1 to 104)
AUTHORS Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chapell, B., Chissoe, S., Dietrich, N., Dubuque, T., Favello, A., Gish, W., Hawkins, E., Hultman, M., Kucaba, T., Lacy, M., Le, M., Le, N., Madis, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L., Rohlfing, T., Scheilenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J., Trevasakis, E., Underwood, K., Wohlmann, P., Waterston, R., Wilson, R. and Marra, M.

JOURNAL Generation and analysis of 280,000 human expressed sequence tags
MEDLINE Genome Res. 6 (9), 807-828 (1996)
COMMENT On Sep 12, 1996 this sequence version replaced gi:1404721.
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -28m3 rev2 from Amersham
High quality sequence stop: 92.

FEATURES
Location/Qualifiers
1..104
/organism="Homo sapiens"
/db_xref="GDB:4594175"
/db_xref="taxon:9606"
/clone="IMAGE:565111"
/clone_lib="Stratagene lung carcinoma 937218"
/tissue_type="lung carcinoma"
/cell_line="NCI-H69"
/dev_stage="cell line NCI-H69"
/lab_host="SOLR (kanamycin resistant)"
/note="Organ: lung; Vector: pBluescript SK-; Site: 1; EcoRI; Site: 2; XhoI; Cloned unidirectionally. Primer: Oligo dt. Small cell carcinoma cell line NCI-H69. Average insert size: 1.0 kb; Uni-ZAP XR Vector: -5' adaptor sequence: 5' GAATTCGGCAGAG 3' -3' adaptor sequence: 5' CTCGAGTTTCTTTTCTTTT 3'."

BASE COUNT 15 a 28 c 26 g 35 t

ORIGIN

Query Match 0.1%; Score 41; DB 29; Length 104;
Best Local Similarity 100.0%; Pred. No. 0.0038;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 676 TGATCTGGCTCAGTCAACCTCCGCTCTGCGTTCAAG 716
|||||
Db 50 TGATCTGGCTCAGTCAACCTCCGCTCTGCGTTCAAG 90

RESULT 11
LOCUS AA669571 95 bp mRNA EST 20-NOV-1997
DEFINITION acled10.s1 Stratagene ovary (#937217) Homo sapiens CDNA clone
IMAGE:856819 3' similar to contains Alu repetitive element; mRNA
sequence.
AA669571
AA669571.1 GI:2631070
EST.
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 95)
AUTHORS Hillier, L., Allen, M., Bowles, L., Dubuque, T., Gelsel, G., Jost, S., Krieman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J., Moore, B., Scheilenberg, K., Steptoe, M., Tan, F., Theising, B., White, Y., Wylie, T., Waterston, R. and Wilson, R.

TITLE WashU-NCI human EST Project
JOURNAL Unpublished (1997)
COMMENT On Nov 6, 1997 this sequence version replaced gi:932677.

CONTACT: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -40m3 fwd. ET from Amersham.

FEATURES
Location/Qualifiers
1..95
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:856819"
/clone_lib="Stratagene ovary (#937217)"
/sex="female"
/dev_stage="49 year old"
/lab_host="SOLR cells (kanamycin resistant)"
/note="Organ: ovary; Vector: Bluescript SK; Site: 1; EcoRI; Site: 2; XhoI; Cloned unidirectionally. Primer: Oligo dt. Total ovary tissue, normal, caucasian. Average insert size: 0.8 kb; Uni-ZAP XR Vector: -5' adaptor sequence: 5' GAATTCGGCAGAG 3' -3' adaptor sequence: 5' CTCGAGTTTCTTTTCTTTT 3'."

BASE COUNT 15 a 30 c 24 g 26 t

ORIGIN

Query Match 0.1%; Score 40; DB 37; Length 95;
Best Local Similarity 100.0%; Pred. No. 0.0084;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 676 TGATCTGGCTCAGTCAACCTCCGCTCTGCGTTCAAG 715
|||||
Db 30 TGATCTGGCTCAGTCAACCTCCGCTCTGCGTTCAAG 69

RESULT 12
LOCUS R67088 96 bp mRNA EST 30-MAY-1995
DEFINITION y130h05.s1 Soares placenta Nb2HP Homo sapiens CDNA clone
IMAGE:140793 3', mRNA sequence.
R67088
R67088.1 GI:839726
EST.
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 96)
AUTHORS Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevasakis, E., Waterston, R., Williamson, A., Wohlmann, P. and Wilson, R.

TITLE The WashU-Merck EST Project
JOURNAL Unpublished (1995)
COMMENT On Apr 5, 1995 this sequence version replaced gi:760768.
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Insert Size: 1093
High quality sequence stops: 65
Source: IMAGE Consortium, LNL
This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Insert length: 1093 Std Error: 0.00
Seq primer: Promega -21m3
High quality sequence stop: 65.
Location/Qualifiers

KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 60)
AUTHORS de la Monte,S. and Wands,J.R.
TITLE Neutral thread protein gene expression and detection of Alzheimer's disease
JOURNAL Patent: US 5830670-A 57 03-NOV-1998;
FEATURES Location/Qualifiers
source 1..60
/organism="unknown"
BASE COUNT 12 a 14 c 15 g 19 t
ORIGIN

Query Match 0.2%; Score 46; DB 5; Length 60;
Best Local Similarity 100.0%; Pred. No. 1e-11;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 25363 GCTAATTTTGTATTTTGTAGACAGAGGTTTCACCGTGTGC 25408
Db 11 GCTAATTTTGTATTTTGTAGACAGAGGTTTCACCGTGTGC 56

RESULT 2
AF087511/c 66 bp mRNA PRI 13-SEP-1999
LOCUS Homo sapiens clone ENAC+22 epithelial sodium channel alpha subunit
DEFINITION (SCNN1A) mRNA, alternatively spliced, partial sequence.
ACCESSION AF087511
VERSION AF087511.1 GI:5870626
KEYWORDS human.
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 66)
AUTHORS Oh,Y. and Warnock,D.G.
TITLE An Alu cassette in the human epithelial sodium channel
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 66)
AUTHORS Oh,Y. and Warnock,D.G.
TITLE Direct Submission
JOURNAL Submitted (26-AUG-1998) Medicine, UAB, 1720 7th Ave. So.,
Birmingham, AL 35294, USA
FEATURES Location/Qualifiers
source 1..66
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="12"
/clone="ENAC+22"
misc-feature 1..>66
/gene="SCNN1A"
/note="alternatively spliced epithelial sodium channel
alpha subunit containing intronic Alu sequence"
gene 1..>66
/gene="SCNN1A"
BASE COUNT 12 a 20 c 24 g 10 t
ORIGIN

Query Match 0.2%; Score 46; DB 40; Length 66;
Best Local Similarity 100.0%; Pred. No. 1e-11;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 8738 CTCGTATCCGCCCGCTCGGCTCCCAAGTCTGGATTACAG 8783
Db 66 CTCGTATCCGCCCGCTCGGCTCCCAAGTCTGGATTACAG 21

RESULT 3
AR051499

LOCUS AR051499 76 bp DNA PAT 29-SEP-1999
DEFINITION Sequence 69 from patent US 5830670.
ACCESSION AR051499
VERSION AR051499.1 GI:5974863
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 76)
AUTHORS de la Monte,S. and Wands,J.R.
TITLE Neutral thread protein gene expression and detection of Alzheimer's disease
JOURNAL Patent: US 5830670-A 69 03-NOV-1998;
FEATURES Location/Qualifiers
source 1..76
/organism="unknown"
BASE COUNT 15 a 19 c 18 g 24 t
ORIGIN

Query Match 0.2%; Score 46; DB 5; Length 76;
Best Local Similarity 100.0%; Pred. No. 1e-11;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 25363 GCTAATTTTGTATTTTGTAGACAGAGGTTTCACCGTGTGC 25408
Db 11 GCTAATTTTGTATTTTGTAGACAGAGGTTTCACCGTGTGC 56

RESULT 4
HUMBRKFAA 80 bp DNA PRI 27-APR-1993
LOCUS Human alpha-galactosidase breakpoint region.
DEFINITION M36131
ACCESSION M36131.1 GI:179541
VERSION M36131.1 GI:179541
KEYWORDS Fabry disease; breakpoint junction; glycosphingolipid catabolism.
SOURCE Human DNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 80)
AUTHORS Kornreich,R., Bishop,D.F. and Desnick,R.J.
TITLE Alpha-galactosidase A gene rearrangements causing Fabry disease:
identification of short direct repeats at breakpoints in an
Alu-rich gene
JOURNAL J. Biol. Chem. 265, 9319-9326 (1990)
MEDLINE 90264427
FEATURES Location/Qualifiers
source 1..80
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 19 a 28 c 20 g 13 t
ORIGIN

Query Match 0.2%; Score 45; DB 9; Length 80;
Best Local Similarity 100.0%; Pred. No. 3.2e-11;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19010 CCCACCTTGCGCTCCCAAGTCTGGATTACAGCATGAGCCAC 19054
Db 13 CCCACCTTGCGCTCCCAAGTCTGGATTACAGCATGAGCCAC 57

RESULT 5
HUMBRKFAA 80 bp DNA PRI 27-APR-1993
LOCUS Human alpha-galactosidase breakpoint family F.
DEFINITION M36132
ACCESSION M36132.1 GI:179542
VERSION M36132.1 GI:179542
KEYWORDS Fabry disease; breakpoint junction; glycosphingolipid catabolism.
SOURCE Human DNA.
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominoidea; Homo.
1 (bases 1 to 80)
Kornreich, R., Bishop, D.F. and Desnick, R.J.
Alpha-galactosidase A gene rearrangements causing Fabry disease:
identification of short direct repeats at breakpoints in an
Alu-rich gene
J. Biol. Chem. 265, 9319-9326 (1990)
MEDLINE
90264427
FEATURES
source
1. 80
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 17 a 29 c 16 g 18 t
ORIGIN

Query Match 0.2%; Score 45; DB 9; Length 80;
Best Local Similarity 100.0%; Pred. No. 3.2e-11;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19010 CCCACCTGGCTGCCAAGTCTGGATACAGCATGAGCCAC 19054
|||||
Db 13 CCCACCTGGCTGCCAAGTCTGGATACAGCATGAGCCAC 57
|||||

RESULT 6
HUMUT578A/c 69 bp DNA STS 19-JUL-1995
LOCUS HUMUT578A 69 bp DNA STS 19-JUL-1995
DEFINITION Human STS UT578, 5' primer bind, sequence tagged site.
ACCESSION U39139 L18336
VERSION L39139.1 GI:642101
KEYWORDS STS; PCR primer; STS sequence; microsatellite DNA; microsatellite
marker; sequence tagged site; tetranucleotide repeat.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominoidea; Homo.
1 (bases 1 to 69)
Gerken, S.C., Matsunami, N., Plaetke, R., Albertsen, H., Ballard, L.,
Wells, R., Lawrence, E., Moore, M., Holik, P.R., Carlson, M., Zhao, X.,
Robertson, M., Bradley, P., Elsnier, T., Tingey, A., Lalouel, J.-M. and
White, R.
Genetic and physical mapping of simple sequence repeat containing
sequence tagged sites from the human genome
Unpublished (1994)
Submitted by: Utah Center for Human Genome Research University of
Utah, Dept. of Human Genetics
2160 Eccles Institute of Human Genetics
Salt Lake City, UT 84112
e-mail: stscorona.med.utah.edu
Primer A: AGTGGAGACCAAGCTGGC
Primer B: AGTGGAGCAAAATCGCATC
End to Label: Primer A
PCR Profile:
Initial Denaturation: 94C 30sec
Cycles Denaturation Annealing Extension 5 94
C 10 sec. 62 C 10 sec. 72 C 20 sec. 30
58 C 10 sec. 72 C 20 sec. Mg++: 1.50 mM
Gel: Acrylamide 7%, Formamide 32%, Urea 34%
Alleles: 1.
FEATURES
source
1. 69
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="19"
14..32
/note="STS UT578 5' end"
/evidence=experimental
BASE COUNT 26 a 16 c 14 g 13 t
ORIGIN

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominoidea; Homo.
1 (bases 1 to 80)
Kornreich, R., Bishop, D.F. and Desnick, R.J.
Alpha-galactosidase A gene rearrangements causing Fabry disease:
identification of short direct repeats at breakpoints in an
Alu-rich gene
J. Biol. Chem. 265, 9319-9326 (1990)
MEDLINE
90264427
FEATURES
source
1. 80
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 17 a 29 c 16 g 18 t
ORIGIN

Query Match 0.2%; Score 44; DB 13; Length 69;
Best Local Similarity 100.0%; Pred. No. 1e-10;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21714 TTCACATGTTGGCAGGCTGTCTGCAACTCTGACCTCAACT 21757
|||||
Db 44 TTCACATGTTGGCAGGCTGTCTGCAACTCTGACCTCAACT 1
|||||

RESULT 7
HUMALUANC/c 65 bp DNA PRI 08-OCT-1994
LOCUS HUMALUANC/c 65 bp DNA PRI 08-OCT-1994
DEFINITION Homo sapiens 4000 year old remains from Nekht-ankh Alu repeat 12:2.
ACCESSION L36836
VERSION L36836.1 GI:556194
KEYWORDS Homo sapiens (individual isolate 4000 year old remains from
Nekht-ankh) liver DNA.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominoidea; Homo.
1 (sites)
Paabo, S.
Ancient DNA: extraction, characterization, molecular cloning, and
enzymatic amplification
Proc. Natl. Acad. Sci. U.S.A. 86 (6), 1939-1943 (1989)
MEDLINE
89184542
FEATURES
source
1. 65
/organism="Homo sapiens"
/isolate="4000 year old remains from Nekht-ankh"
/db_xref="taxon:9606"
/tissue_type="liver"
repeat_region
repeat_region /rpl_family="Alu"
BASE COUNT 17 a 17 c 21 g 10 t
ORIGIN

Query Match 0.1%; Score 43; DB 9; Length 65;
Best Local Similarity 100.0%; Pred. No. 3.3e-10;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8735 GACCTGTATCCGCGCTGCTGCCAAGTCTGGAT 8777
|||||
Db 65 GACCTGTATCCGCGCTGCTGCCAAGTCTGGAT 23
|||||

RESULT 8
HUMALUANC/c 67 bp DNA PRI 08-OCT-1994
LOCUS HUMALUANC/c 67 bp DNA PRI 08-OCT-1994
DEFINITION Homo sapiens 4000 year old remains from Nekht-ankh Alu repeat
fragment 12:5.
ACCESSION L36843
VERSION L36843.1 GI:556196
KEYWORDS Alu repeat.
SOURCE Homo sapiens (individual isolate 4000 year old remains from
Nekht-ankh) liver DNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominoidea; Homo.
1 (sites)
Paabo, S.
Ancient DNA: extraction, characterization, molecular cloning, and
enzymatic amplification
Proc. Natl. Acad. Sci. U.S.A. 86 (6), 1939-1943 (1989)
MEDLINE
89184542
FEATURES
source
1. 67
/organism="Homo sapiens"
/isolate="4000 year old remains from Nekht-ankh"
/db_xref="taxon:9606"
/tissue_type="liver"

```
repeat_region <1..>67
/rpt_family="Alu"
BASE COUNT 15 a 18 c 23 g 11 t
ORIGIN

Query Match 0.1%; Score 43; DB 9; Length 67;
Best Local Similarity 100.0%; Pred. No. 3.3e-10;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25448 GCCTCAGCCTCCCAAGTGTGGATTACAGGTGAGCCACC 25490
|||||
50 GCCTCAGCCTCCCAAGTGTGGATTACAGGTGAGCCACC 8

RESULT 9
A68621/c A68621 40 bp DNA PAT 06-MAY-1999
LOCUS Sequence 1 from Patent WO9801573.
DEFINITION A68621
ACCESSION A68621.1 GI:4759648
VERSION
KEYWORDS
SOURCE unidentified.
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 40)
AUTHORS Resnick,M.A., Laitinov,V.L., Koudrina,N.Y. and Perkins,E.L.
TITLE TRANSFORMATION-ASSOCIATED RECOMBINATION CLONING
JOURNAL Patent: WO 9801573-A 15-JAN-1998;
US HEALTH (US)
FEATURES
Source Location/Qualifiers
1..40
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 7 a 12 c 13 g 8 t
ORIGIN

Query Match 0.1%; Score 40; DB 5; Length 40;
Best Local Similarity 100.0%; Pred. No. 1.1e-08;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8755 TCGGCTCCCAAGTGTGGATTACAGCGTGAGCCACC 8794
|||||
40 TCGGCTCCCAAGTGTGGATTACAGCGTGAGCCACC 1

RESULT 10
HUMALCE221/c HUMALCE221 103 bp ss-RNA PRI 15-APR-1994
LOCUS Human carcinoma cell-derived Alu RNA transcript, clone CE221.
DEFINITION M87896
ACCESSION M87896.1 GI:174874
VERSION
KEYWORDS Alu repeat.
SOURCE Homo sapiens male embryo carcinoma cDNA to other RNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 103)
AUTHORS Smett,D., Richer,C., Deragon,J.-M. and Labuda,D.
TITLE Alu RNA transcripts in human embryonal carcinoma cells. Model of
post-transcriptional selection of master sequences
JOURNAL J. Mol. Biol. (1992) In press
FEATURES
Source Location/Qualifiers
1..103
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_line="NTera2D1"
/dev_stage="embryo"
/sex="male"
/tissue_type="carcinoma"
BASE COUNT 25 a 27 c 33 g 18 t
ORIGIN
```

```
Query Match 0.1%; Score 40; DB 9; Length 103;
Best Local Similarity 100.0%; Pred. No. 1e-08;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18828 CGATCTGGCTCACTGCACCTCCGCTCCGGGTTCAAG 18867
|||||
87 CGATCTGGCTCACTGCACCTCCGCTCCGGGTTCAAG 48

RESULT 11
AR051522 AR051522 85 bp DNA PAT 29-SEP-1999
LOCUS Sequence 92 from patent US 5830670.
DEFINITION AR051522
ACCESSION AR051522.1 GI:5974886
VERSION
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 85)
AUTHORS de la Monte,S. and Wands,J.R.
TITLE Neural thread protein gene expression and detection of Alzheimer's
disease
JOURNAL Patent: US 5830670-A 92 03-NOV-1998;
FEATURES
Source Location/Qualifiers
1..85
/organism="unknown"
BASE COUNT 13 a 27 c 25 g 20 t
ORIGIN

Query Match 0.1%; Score 39; DB 5; Length 85;
Best Local Similarity 100.0%; Pred. No. 3.3e-08;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17462 GCCAGATGCTCGATCTCCTGACCTCGTGATCCGCC 17500
|||||
9 GCCAGATGCTCGATCTCCTGACCTCGTGATCCGCC 47

RESULT 12
HUMBRKFAC HUMBRKFAC 80 bp DNA PRI 27-APR-1993
LOCUS Human alpha-galactosidase breakpoint region.
DEFINITION M36133
ACCESSION M36133.1 GI:179543
VERSION
KEYWORDS Fabry disease; breakpoint junction; glycosphingolipid catabolism.
SOURCE Human DNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 80)
AUTHORS Kornreich,R., Bishop,D.F. and Desnick,R.J.
TITLE Alpha-galactosidase A gene rearrangements causing Fabry disease:
Identification of short direct repeats at breakpoints in an
Alu-rich gene
JOURNAL J. Biol. Chem. 265, 9319-9326 (1990)
FEATURES
MEDLINE 90264427
SOURCE Location/Qualifiers
1..80
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 18 a 28 c 17 g 17 t
ORIGIN

Query Match 0.1%; Score 37; DB 9; Length 80;
Best Local Similarity 100.0%; Pred. No. 3.4e-07;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19018 GGCTCCCAAGTGTGGATTACAGGATGAGCCAC 19054
```

```
|||||
Db      21 GGCCTCCCAAAGTCTGGGATTACAGCATGACCCAC 57
RESULT  13
LOCUS   HUMDLRL 90 bp DNA PRI 11-JAN-1995
DEFINITION Human low density lipoprotein receptor Intron A Alu repeat.
ACCESSION K03555.1 GI:187104
VERSION   K03555.1 GI:187104
KEYWORDS  Alu repeat; low density lipoprotein receptor-1; repeat region.
SOURCE    Human DNA.
ORGANISM  Homo sapiens
REFERENCE 1 (bases 1 to 90)
AUTHORS   Eutheria; Primates; Carnivora; Mammalia;
           Lehman,M.A., Goldstein,J.L., Russell,D.W. and Brown,M.S.
           Duplication of seven exons in LDL receptor gene caused by Alu-Alu
           recombination in a subject with familial hypercholesterolemia
           Cell 48 (5), 827-835 (1987)
JOURNAL   87131094
COMMENT   Clean copy of sequence kindly provided by M. Lehman (22-APR-1987).
FEATURES  Location/Qualifiers
           source          1..90
                           /organism="Homo sapiens"
                           /db_xref="taxon:9606"
                           /map="19p13.3"
misc_feature 1
              /gene="LDLR"
              /note="hgml: G00-119-362"
              /note="hgml: 24 g 19 t
BASE COUNT  17 a 30 c 24 g 19 t
ORIGIN      Chromosome 19p13.2-p13.1.
```

```
Query Match 0.1%; Score 37; DB 9; Length 90;
Best Local Similarity 100.0%; Pred. No. 3.4e-07;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

OY 19018 GGCCTCCCAAAGTCTGGGATTACAGCATGACCCAC 19054

Db 42 GGCCTCCCAAAGTCTGGGATTACAGCATGACCCAC 78

```
RESULT 14
LOCUS   I31263 40 bp DNA PAT 06-FEB-1997
DEFINITION Sequence 175 from patent US 5582979.
ACCESSION I31263
VERSION   I31263.1 GI:1822054
KEYWORDS
SOURCE    Unknown.
ORGANISM  Unknown.
REFERENCE 1 (bases 1 to 40)
AUTHORS   Weber,J.L.
           Length polymorphisms in (dc-da).sub.n.(dc-dt).sub.n sequences and
           method of using the same
           Patent: US 5582979-A 175 10-DEC-1996;
           Location/Qualifiers
           source          1..40
                           /organism="unknown"
BASE COUNT  20 a 19 c 0 g 1 t
ORIGIN
```

```
Query Match 0.1%; Score 36; DB 5; Length 40;
Best Local Similarity 100.0%; Pred. No. 1.1e-06;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

OY 28999 TGTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 29034

Db 36 TGTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 1

```
RESULT 15
LOCUS   I31159 46 bp DNA PAT 06-FEB-1997
DEFINITION Sequence 71 from patent US 5582979.
ACCESSION I31159
VERSION   I31159.1 GI:1821950
KEYWORDS
SOURCE    Unknown.
ORGANISM  Unknown.
REFERENCE 1 (bases 1 to 46)
AUTHORS   Weber,J.L.
           Length polymorphisms in (dc-da).sub.n.(dc-dt).sub.n sequences and
           method of using the same
           Patent: US 5582979-A 71 10-DEC-1996;
           Location/Qualifiers
           source          1..46
                           /organism="unknown"
BASE COUNT  23 a 22 c 0 g 1 t
ORIGIN
```

```
Query Match 0.1%; Score 36; DB 5; Length 46;
Best Local Similarity 100.0%; Pred. No. 1.1e-06;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

OY 28999 TGTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 29034

Db 44 TGTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 9

Search completed: April 1, 2000, 10:12:16
Job time: 283685 sec

PR 21-APR-1989; US-341562.
 PR 05-SEP-1991; US-754351.
 PR 04-APR-1994; US-222177.
 PA (MARS-) MARSHFIELD CLINIC.
 PI Weber J.L.
 DR WPI: 97-042299/04.
 PT Detection of polymorphic genetic markers of the form
 PT (dc-da)n(dg-dt)n - using novel nucleic acid mols. as primers
 PS Disclosure; Column 9-10; 186pp; English.
 CC The invention relates to the isolation of polymorphic repeat sequences
 CC having the sequence (dc-da)n.(dg-dt)n which can be used as genetic
 CC markers. Primers based on these sequences can be used to detect these
 CC repeats, especially for use in e.g. paternity or maternity testing,
 CC human genetic analysis such as linkage analysis of genetic disease,
 CC commercial animal or plant breeding or pedigree analysis. Clones
 CC containing the repeat sequences were isolated by hybridization of
 CC chromosome-specific phage libraries with a synthetic poly(dc-da).(dg-dt)
 CC probe. Over 100 repeat blocks were isolated. The inserts from the clones
 CC were amplified by primers T65798-T66047. Those clones where the repeat
 CC sequence has been determined are shown in T65704-T6797. This repeat
 CC sequence is from the marker clone Mdf42 which contains the repeat
 CC sequence having the formula: (CA)16T(AC)3.5.
 Sequence 40 Bp; 20 A; 19 C; 0 G; 1 T;

				0.1%;	Score 36;	DB 1;	Length 40;
				Best Local Similarity	100.0%;	Pred. No.	2.7;
Matches	36;	Conservative	0;	Mismatches	0;	Indels	0; Gaps 0;
QY	28999	TGATGTCTGTCTGTCTGTCTGTCTGTCTGC	29034				
Dz	36						
		TGAATGTGTGTGTGTGTGTGTGTGTGTG	1				

RESULT 3
 ID T65709/c
 AC T65709 standard; DNA: 46 BP.
 DT 17-JUN-1997 (first entry)
 DE Repeat sequence from polymorphic marker clone Mfd7.
 KW Polymorphism; repeat sequence; genetic marker; primer; amplification;
 KW PCR; polymerase chain reaction; paternity; maternity; human; pedigree;
 KW linkage analysis; genetic disease; animal; plant; breeding; locus;
 KW hybridisation; chromosome; ds.
 OS Homo sapiens.
 PN US5582979-A.
 PD 10-DEC-1996.
 PF 21-APR-1989; 341562.
 PR 21-APR-1989; US-341562.
 PR 05-SEP-1991; US-754351.
 PR 04-APR-1994; US-222177.
 PA (MARS-) MARSHFIELD CLINIC.
 PI Weber JL;
 DR WPI: 97-042299/04..
 PT Detection of polymorphic genetic markers of the form
 PT (dc-da)n(dg-dt)n - using novel nucleic acid mols. as primers
 PS Claim 1; Column 9-10; 186pp; English.
 CC The invention relates to the isolation of polymorphic repeat sequences
 CC having the sequence (dc-da)n.(dg-dt)n which can be used as genetic
 CC markers. Primers based on these sequences can be used to detect these
 CC repeats, especially for use in e.g. paternity or maternity testing,
 CC human genetic analysis such as linkage analysis of genetic disease,
 CC commercial animal or plant breeding or pedigree analysis. Clones
 CC containing the repeat sequences were isolated by hybridisation of
 CC chromosome-specific phage libraries with a synthetic poly(dc-da).(dg-dt)
 CC probe. Over 100 repeat blocks were isolated. The inserts from the clones
 CC were amplified by primers T65798-T66047. Those clones where the repeat
 CC sequence has been determined are shown in T65704-797. This repeat
 CC sequence is from the marker clone Mdf7 which contains the repeat
 CC sequence having the formula: (CA)20TA(CA)2.
 SQ Sequence 46 BP; 23 A; 22 C; 0 G; 1 T;

Query Match	0.1%;	Score 36;	DB 1;	Length 46;
Best Local Similarity	100.0%;	Pred. No. 2 6;		
Matches 36;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

Qy 28999 TGTATGCTGCTGCTGCTGCTGCTGCTGCTG 29034
|||||
Db 44 TGTATGCTGCTGCTGCTGCTGCTGCTGCTG 9

RESULT 4

AC Q34122;

DE Sequence of a microsatellite
PCB: selection: primers: OPT

OS Bos taurus.

PD 00-AUG-1992;
PE 15-JAN-1992; U00340.

PA (GEM-) GENMARK.
PT George M. Massay, TM.

PT Polymorphic bovine DNA marker

The sequence is that of a bcc

CC 250 and 500 bp with an (AC)₁₀ repeat

(T6)n > 9 microsatellites in

is summarised in the specified sequences upstream and downstream of the gene.

CC amplification of the correspond-

CC for parentage testing, and
CC loci, or genes involved the

CC See also Q33501-34437.
SO Sequence 49 BP: 7 A:

1

Best Local Similarity 100.00
Matches 36; Conservative

[illegible]

3
RESULTS
Q33870

AC Q33870; 03-FEB-1993 (44 sec + 0.4 sec)

KW PCR; selection; primers; OPT

OS Bos taurus.
PN W09213102-A

PF 15-JAN-1992; U00340.

PA (GENM-) GENMARK.
PI Georges M. Massey JM:

PT Polymorphic bovine DNA marker

PT Polymorphic bovine DNA markers - used in genetic identification,
 PT gene mapping, and selective breeding
 PS Table 7: Page 379: 517pp: English.
 CC The sequence is that of a bovine microsatellite sequence obtd.
 CC by screening a library of bovine MboI DNA fragments of between
 CC 250 and 500 bp with an (AC)₁₅ and a (TC)₁₅ oligonucleotide probe.
 CC One out of 50 clones cross-hybridised. Assuming independent
 CC distribution of microsatellites and MboI sites, the frequency of
 CC (TC)_n >9 microsatellites in the bovine genome is estimated at >100,
 CC 000. The sequence information for ca. 230 such bovine microsatellites
 CC is summarised in the specification and indexed herein (see below).
 CC The sequences upstream and downstream of the microsatellite sequence
 CC were used to generate the required PCR primers for *in vitro*
 CC amplification of the corresp. microsatellite (using the program
 CC OmpPrm). The microsatellites may be used to identify individuals,
 CC for parentage testing, and in the genetic mapping of economically
 CC local, or genes involved the determinism of economically important
 CC traits esp. in cattle, to allow selective breeding.
 CC See also Q33501-34437.
 Q3 Sequence 49 BP; 7 A; 0 C; 17 G; 25 T;

```
Query Match          0.1%; Score 36; DB 1; Length 49;
Best Local Similarity 100.0%; Pred. No. 2.5;
Matches   36; Conservative    0; Mismatches      0; Indels      0; Gaps      0;
```

RESULT	5
033870	
ID	Q33870 standard; DNA; 54 BP.
AC	Q33870.
DT	02-FEB-1993 (first entry)
DE	Microsatellite sequence from clone TGLA272.
KW	PCR; selection; primers; OTFPRIM; breeding; cattle; parentage;
KW	genetic mapping; traits; amplification; ss.
OS	Bos taurus.
PN	W09213102-A.
PD	06-AUG-1992.
PF	15-JAN-1992; U00340.
PR	15-JAN-1991; US-642342.
PA	(GENM-) GENMARK.
PI	Georges M. Massey JM;
DR	WPI; 92-284684/34.
PT	Polymorphic bovine DNA markers - used in genetic identification,

PT gene mapping, and selective breeding
PS Table 7, page 277, 517pp; English.
CC The sequence is that of a bovine microsatellite sequence obt'd. by
CC screening a library of bovine MboI DNA fragments of between
CC 250 and 500 bp with an (AC)₁₅ and a (TC)₁₅ oligonucleotide probe.
CC One out of 50 clones cross-hybridised. Assuming independent
CC distribution of microsatellites and MboI sites, the frequency of
CC (T6)n >9 microsatellites in the bovine genome is estimated at >100,
CC 000. The sequence information for ca. 230 such bovine microsatellites
CC is summarised in the specification and indexed herein (see below).
CC The sequences upstream and downstream of the microsatellite sequence
CC were used to generate the required PCR primers for in vitro
CC amplification of the corresp. microsatellite (using the program
CC OPRPRM). The microsatellites may be used to identify individuals,
CC for parentage testing, and in the genetic mapping of economic trait
CC loci, or genes involved in the determinism of economically important
CC traits esp. in cattle, to allow selective breeding.
CC See also Q33501-34437.
SO Sequence 54 Bp; 1 A; 0 C; 26 G; 27 T;

```
Query Match      0.1%; Score 36; DB 1; Length 54;
Best Local Similarity 100.0%; Pred. NO. 2.5;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY	28939	TCGATGCTGCTGCTGCTGCTGCTGCTG	29034
DB	11	TCGATGCTGCTGCTGCTGCTGCTGCTG	46

RESULT	6
Q27391	
ID	Q27391 standard; DNA; 35 BP.

DE Inter-Au specific primer PDJ3.
KW Polymerase chain reaction; PCR; repetitive element; ss.
OS Synthetic.

PF 24-JAN-1992: NL0018.
PR 25-JAN-1991: NL-000132.
PA (INGE-) INGENY BV. 1991.

PT detection of genetic variation by 2-D electrophoresis of fragments - and hybridisation with labelled probes, carried out on fragments consisting of Inter-repeat sequences, generated by

CC Primer PDJ33 is one of several primers which are preferred for use
CC in amplifying inter-Alu regions of DNA. The amplified fragments are
CC then subjected to 2-D electrophoresis on the basis of length and

CC be used to detect genetic variation.
CC See Q27389-Q27404 and Q33141-Q33144.
SQ Sequence 35 BP; 8 A; 10 C; 11 G; 6 T;

```

Query Match      0.1%; Score 35; DB 1; Length 35;
Best Local Similarity 100.0%; Pred. No. 4.9;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy	8758	GCCTCCCAAGTGTGGATTACAGCGGTGAGCCA	8792
Db	1	GCCTCCCAAGTGCTGGATTACAGGCGGTGAGCCA	35

T65737/C
ID T65737 standard; DNA; 45 BP.
AC T65737;

DT 17-JUN-1997 (first entry)
DE Repeat sequence from polymorphic marker clone Mf36.
KW Polymorphism; repeat sequence; genetic marker; primer; amplification;
KW PCR; polymerase chain reaction; paternity; maternity; human; pedigree;
KW linkage analysis; genetic disease; animal; plant; breeding; locus;
KW hybridisation; chromosome; ds.

PD	10-DEC-1996.	341562.
PF	21-APR-1989.	US-341562.
PR	21-APR-1989.	US-341562.

PA (MARS-) MARSHFIELD CLINIC.
PI Weber TL;
DR WPI; 97-042299/04.

PS disclosure; Column 9-10; 186pp; English.
CC The invention relates to the isolation of polymorphic repeat sequences
CC having the sequence (dc-dan).(dc-dtn) which can be used as genetic
CC

[illegible]

CC sequence having the formula: (AC)¹⁵AT/ACG15A.
CC sequence is from the marker clone Mdc36 which contains the repeat
CC sequence has been determined are shown in T65704-797. This repeat
CC sequence was amplified by primers T65798-T66047. Those clones where the repeat
CC sequence is from the marker clone Mdc36 which contains the repeat

```
Query Match      0.1%  Score 35;  DB 1;  Length 45;
Best Local Similarity 100.0%  Pred. No. 4 4;
```

[illegible]

RESULT 8
O34053/C
ID O34053 standard: DNA: 62 BP.

D1 02-Feb-1993 (first entry)
 DE Microsatellite sequence from clone TGLA48.
 KR PCR; selection: primers; OPTIPRIM; breeding: cattle; parentage:
 KW genetic mapping: traits: amplification: ss.
 KM

FN
MO3213-02-A.
PD
06-AUG-1992.
PF
15-JAN-1992: U00340.
R
15-JAN-1991: US-642342.
R

PT Polymorphic bovine DNA markers - used in genetic identification,
gene mapping, and selective breeding

screening a library of bovine MBOI DNA fragments of between 250 and 500 bp with an (AC)₁₅ and a (TC)₁₅ oligonucleotide probe. One out of 50 clones cross-hybridised. Assuming independent

CC 000. The sequence information for ca. 230 such bovine microsatellites
CC is summarised in the specification and indexed herein (see below).
CC The sequences upstream and downstream of the microsatellite sequence
CC

CC were used to generate the regulated PCR primers for *in vitro* CC

CC The sequences upstream and downstream of the microsatellite sequence
CC were used to generate the required PCR primers for *in vitro*
CC amplification of the corresp. microsatellite (using the program
CC OPRPRM). The microsatellites may be used to identify individuals,
CC for parentage testing, and in the genetic mapping of economic trait
CC loci, or genes involved in the determination of economically important
CC traits esp. in cattle, to allow selective breeding.
CC See also Q33501-34437.
SQ Sequence 56 BP; 5 A; 0 C; 23 G; 28 T;

```
Query Match      0.1%; Score 34; DB 1; Length 56;
Best Local Similarity 100.0%; Pred. No. 7;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Oy 29001 TATGTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 29034
      |||||
Db    9 TATGTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 42
```

RESULT	12
Q33936	
ID	Q33936 standard; DNA; 79 BP.

DT 02-FEB-1993 (first entry)
DE Microsatellite sequence from clone TGLA34.
KM PCR, selection; primers: OPTPRIM; breeding; cattle; parentage
KM genetic mapping; traits; amplification; ss.

PD	06-AUG-1992.	U00340.
PF	15-JAN-1992;	US-642342.
PR	15-JAN-1991;	

PT Polymorphic bovine DNA markers - used in genetic identification, gene mapping, and selective breeding

CC screening a library of bovine MOI DNA fragments of between
CC 250 and 500 bp with an (AC)₁₅ and a (TC)₁₅ oligonucleotide probe.
CC One out of 50 clones cross-hybridised. Assuming independent

CC The sequences upstream and downstream of the microsatellite sequence
CC is summarised in the specification and indexed herein (see below).

CC OPTIPRIM). The microsatellites may be used to identify individuals
CC for parentage testing, and in the genetic mapping of economic trait
CC loci, or genes involved the determinism of economically important

Sequence	79 BP;	16 A;	0 C;	24 G;	39 T;
50					

	0.1%;	Score 34;	DB 1;	Length 79;
Query Match				
Best Local Similarity	100.0%;	Pred. No. 6.2;		
Matches 34;	Conservative 0;	Mismatches 0;	Gaps 0.	
29001 TATGTGTTGGTGTGTGTGTGTGTGTG	29034			

RESULT 13

AC	T26728;
DT	22-OCT-1996 (first entry)
DE	Human gene signature HUGS08978.

KW Gene signature; messenger RNA; mRNA; relative abundance; frequency
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.

PD 01-JUN-1995.
PF 11-NOV-1994; J01916.
PR 12-NOV-1993; JP-355504

PI Matsubara K, Okubo K;
DR WPI: 95-206931/27.
PT Identifying gene signatures in 3'-directed human cDNA library - e.g

PT tissues
PS Claim 1; Page 2158; 2245bp; Japanese
CC A single-stranded DNA (or its complemen

CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the

all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library is constructed so as to reflect accurately the relative abundance of different mRNAs in the particular tissue from which it was derived.

CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.
SQ Sequence 97 BP; 19 A; 27 C; 20 G; 28 T;

Query Match	0.1%	Score 34	DB 1	Length 97
Best Local Similarity	100.0%	Pred. No. 5.7		
Matches 34; Conservative	0	Mismatches	0	Gaps 0

Db 25 AGTGTGGATTACAGGCATGAGCCACTGCGCC 58

ID	Q33855	standard; DNA; 44 BP
AC	Q33855;	

KM PCR; selection; primers; OPTIPRM; breeding; cattle; parentage
KM genetic mapping; traits; amplification; ss.
OS Bos taurus.

PF 15-JAN-1992; U00340.
PR 15-JAN-1991; US-642342
PA (GENM-) GENMARK.

PT Polymorphic bovine DNA markers - used in genetic identification, gene mapping, and selective breeding
PS Table 7; Page 271; 517pp; English.

CC 250 and 500 bp with an (AC)₁₅ and a (TC)₁₅ oligonucleotide probe.
CC One out of 50 clones cross-hybridised. Assuming independent
CC distribution of microsatellites and MboI sites, the frequency of

CC is summarised in the specification and indexed herein (see below).
 CC The sequences upstream and downstream of the microsatellite sequence
 CC were used to generate the required PCR primers for in vitro

CC amplification of the corresp. microsatellite (using the program
CC OPIPRIM). The microsatellites may be used to identify individuals,
CC for parentage testing, and in the genetic mapping of economic trait
CC loci, or genes involved the determinism of economically important
CC traits esp. in cattle, to allow selective breeding.
CC See also Q33501-34437.
SQ Sequence 44 BP; 22 A; 0 C; 0 G; 22 T;

Query Match 0.1%; Score 33; DB 1; Length 44;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 13052 ATATATATATATATATATATATATATATA 13084
|||||
DB 3 ATATATATATATATATATATATATATATA 35

RESULT 15

ID Q33855/c
AC Q33855. standard; DNA; 44 BP.

DT 02-FEB-1993 (first entry)
DE Microsatellite sequence from clone TGLA261.
KW PCR; selection; primers; OPIPRIM; breeding; cattle; parentage;
KW genetic mapping; traits; amplification; ss.
OS Bos taurus.

PN W09213102-A.

PD 06-AUG-1992.

PF 15-JAN-1992; U00340.

PR 15-JAN-1991; US-642342.

PA (GENM-) GENMARK.

PI Georges M. Massey JM;

DR WPI: 92-284684/34.

PT Polymorphic bovine DNA markers - used in genetic identification,

PT gene mapping, and selective breeding

PS Table 7; Page 271; 517pp; English.

CC The sequence is that of a bovine microsatellite sequence obtd. by

CC screening a library of bovine MboI DNA fragments of between

CC 250 and 500 bp with an (AC)₁₅ and a (TC)₁₅ oligonucleotide probe.

CC One out of 50 clones cross-hybridised. Assuming independent

CC distribution of microsatellites and MboI sites, the frequency of

CC (T6)_n >9 microsatellites in the bovine genome is estimated at >100,

CC 000. The sequence information for ca. 230 such bovine microsatellites

CC is summarised in the specification and indexed herein (see below).

CC The sequences upstream and downstream of the microsatellite sequence

CC were used to generate the required PCR primers for in vitro

CC amplification of the corresp. microsatellite (using the program

CC OPIPRIM). The microsatellites may be used to identify individuals,

CC for parentage testing, and in the genetic mapping of economic trait

CC loci, or genes involved the determinism of economically important

CC traits esp. in cattle, to allow selective breeding.

CC See also Q33501-34437.

SQ Sequence 44 BP; 22 A; 0 C; 0 G; 22 T;

Query Match 0.1%; Score 33; DB 1; Length 44;

Best Local Similarity 100.0%; Pred. No. 13;

Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 13052 ATATATATATATATATATATATATATATA 13084
|||||
DB 42 ATATATATATATATATATATATATATATA 10

Search completed: April 1, 2000, 12:51:40
Job time: 291698 sec

RESULT 2
US-08-340-426D-57
Sequence 57, Application US/08340426D
Patent No. 5948634
GENERAL INFORMATION:
APPLICANT: de la Monte, Suzanne
TITLE OF INVENTION: Neutral Thread Protein Gene Expression and Detection
TITLE OF INVENTION: of Alzheimer's Disease
NUMBER OF SEQUENCES: 121
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/340,426D
FILING DATE: 14-NOV-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609.3840002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 57:
SEQUENCE CHARACTERISTICS:
LENGTH: 60 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
US-08-340-426D-57
Query Match 0.2%; Score 46; DB 4; Length 60;
Best Local Similarity 100.0%; Pred. No. 6.2e-06;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 25363 GCTAATTTTGTATTTTGTAGACAGAGGTTTCACCGTGTGGC 25408
|||||
DB 11 GCTAATTTTGTATTTTGTATTTTGTAGACAGAGGTTTCACCGTGTGGC 56
RESULT 3
US-08-450-673C-57
Sequence 57, Application US/08450673C
Patent No. 5948888
GENERAL INFORMATION:
APPLICANT: de la Monte, Suzanne
TITLE OF INVENTION: Neutral Thread Protein Gene Expression and Detection
TITLE OF INVENTION: of Alzheimer's Disease
NUMBER OF SEQUENCES: 121
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/450,673C
FILING DATE: 30-MAY-1995
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609.3840004
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 57:
SEQUENCE CHARACTERISTICS:
LENGTH: 60 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
US-08-450-673C-57
Query Match 0.2%; Score 46; DB 4; Length 60;
Best Local Similarity 100.0%; Pred. No. 6.2e-06;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 25363 GCTAATTTTGTATTTTGTAGACAGAGGTTTCACCGTGTGGC 25408
|||||
DB 11 GCTAATTTTGTATTTTGTATTTTGTAGACAGAGGTTTCACCGTGTGGC 56
RESULT 4
PCT-US95-17111A-57
Sequence 57, Application PC/TUS951711A
GENERAL INFORMATION:
APPLICANT: de la Monte, Suzanne
TITLE OF INVENTION: Neutral Thread Protein Gene Expression and
TITLE OF INVENTION: Detection of Alzheimer's Disease
NUMBER OF SEQUENCES: 121
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/17111A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/340,426
FILING DATE: 14-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609.3840002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 57:
SEQUENCE CHARACTERISTICS:
LENGTH: 60 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
PCT-US95-17111A-57

Query Match 0.2%; Score 46; DB 6; Length 60;
Best Local Similarity 100.0%; Pred. No. 6.2e-06;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25363 GCTAATTTTGTATTTTGTAGACAGGAGGTTTCACCGTGTGGC 25408
|||||
DB 11 GCTAATTTTGTATTTTGTAGACAGGAGGTTTCACCGTGTGGC 56

RESULT 5
US-08-454-557C-69
Sequence 69, Application US/08454557C
Patent No. 5830670
GENERAL INFORMATION:
APPLICANT: de la Monte, Suzanne
TITLE OF INVENTION: Neutral Thread Protein Gene Expression and Detection
TITLE OF INVENTION: of Alzheimer's Disease
NUMBER OF SEQUENCES: 121
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/454,557C
FILING DATE: 30-MAY-1995
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609.3840003
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 69:
SEQUENCE CHARACTERISTICS:
LENGTH: 76 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
US-08-454-557C-69

Query Match 0.2%; Score 46; DB 3; Length 76;
Best Local Similarity 100.0%; Pred. No. 5.9e-06;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25363 GCTAATTTTGTATTTTGTAGACAGGAGGTTTCACCGTGTGGC 25408
|||||
DB 11 GCTAATTTTGTATTTTGTAGACAGGAGGTTTCACCGTGTGGC 56

RESULT 6
US-08-340-426D-69
Sequence 69, Application US/08340426D
Patent No. 5948634
GENERAL INFORMATION:
APPLICANT: de la Monte, Suzanne
TITLE OF INVENTION: Neutral Thread Protein Gene Expression and Detection
TITLE OF INVENTION: of Alzheimer's Disease
NUMBER OF SEQUENCES: 121
CORRESPONDENCE ADDRESS:

ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/340,426D
FILING DATE: 14-NOV-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609.3840002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 69:
SEQUENCE CHARACTERISTICS:
LENGTH: 76 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
US-08-340-426D-69

Query Match 0.2%; Score 46; DB 4; Length 76;
Best Local Similarity 100.0%; Pred. No. 5.9e-06;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25363 GCTAATTTTGTATTTTGTAGACAGGAGGTTTCACCGTGTGGC 25408
|||||
DB 11 GCTAATTTTGTATTTTGTAGACAGGAGGTTTCACCGTGTGGC 56

RESULT 7
US-08-450-673C-69
Sequence 69, Application US/08450673C
Patent No. 5948888
GENERAL INFORMATION:
APPLICANT: de la Monte, Suzanne
TITLE OF INVENTION: Neutral Thread Protein Gene Expression and Detection
TITLE OF INVENTION: of Alzheimer's Disease
NUMBER OF SEQUENCES: 121
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/450,673C
FILING DATE: 30-MAY-1995
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609.3840004
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540

INFORMATION FOR SEQ ID NO: 69:
SEQUENCE CHARACTERISTICS:
LENGTH: 76 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
US-08-450-673C-69

Query Match 0.2%; Score 46; DB 4; Length 76;
Best Local Similarity 100.0%; Pred. No. 5.9e-06;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 11 GCTAATTTTGTATTTTGTAGACAGACAGGTTTCACCGTGTGGC 25408
|||||

RESULT 8
PCT-US95-17111A-69
Sequence 69, Application PC/TUS9517111A

GENERAL INFORMATION:
APPLICANT: de la Monte, Suzanne
TITLE OF INVENTION: Neural Thread Protein Gene Expression and
TITLE OF INVENTION: Detection of Alzheimer's Disease
NUMBER OF SEQUENCES: 121
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/17111A
CLASSIFICATION:
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/340,426
FILING DATE: 14-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609.3840002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2540
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 69:
SEQUENCE CHARACTERISTICS:
LENGTH: 76 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
PCT-US95-17111A-69

Query Match 0.2%; Score 46; DB 6; Length 76;
Best Local Similarity 100.0%; Pred. No. 5.9e-06;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 25363 GCTAATTTTGTATTTTGTAGACAGACAGGTTTCACCGTGTGGC 25408
|||||

RESULT 9
US-08-454-557C-92

Sequence 92, Application US/08454557C
Patent No. 5830670
GENERAL INFORMATION:
APPLICANT: de la Monte, Suzanne
TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection
TITLE OF INVENTION: of Alzheimer's Disease
NUMBER OF SEQUENCES: 121
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/454,557C
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609.3840003
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2540
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 92:
SEQUENCE CHARACTERISTICS:
LENGTH: 85 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
US-08-454-557C-92

Query Match 0.1%; Score 39; DB 3; Length 85;
Best Local Similarity 100.0%; Pred. No. 0.00086;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 17462 GCCAGATGCTGTCGATCCTGACCTCGTATCGGCC 17500
|||||

RESULT 10
US-08-340-426D-92
Sequence 92, Application US/08340426D
Patent No. 5948634
GENERAL INFORMATION:
APPLICANT: de la Monte, Suzanne
TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection
TITLE OF INVENTION: of Alzheimer's Disease
NUMBER OF SEQUENCES: 121
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/340,426D

Query Match 0.1%; Score 39; DB 3; Length 85;
Best Local Similarity 100.0%; Pred. No. 0.00086;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 9 GCCAGATGCTGTCGATCCTGACCTCGTATCGGCC 47
|||||

FILING DATE: 14-NOV-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609.3840002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 92:
SEQUENCE CHARACTERISTICS:
LENGTH: 85 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
US-08-340-426D-92

Query Match 0.1%; Score 39; DB 4; Length 85;
Best Local Similarity 100.0%; Pred. No. 0.00086;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17462 GCCAGATGGTCTGATCTCTGACCTGATCGGCC 17500
|||||
DB 9 GCCAGATGGTCTGATCTCTGACCTGATCGGCC 47

RESULT 11
US-08-450-673C-92
Sequence 92, Application US/08450673C
Patent No. 594888
GENERAL INFORMATION:
APPLICANT: de la Monte, Suzanne
APPLICANT: Mands, Jack R.
TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection
NUMBER OF SEQUENCES: 121
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/450.673C
FILING DATE: 30-MAY-1995
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609.3840004
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 92:
SEQUENCE CHARACTERISTICS:
LENGTH: 85 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
US-08-450-673C-92

Query Match 0.1%; Score 39; DB 4; Length 85;
Best Local Similarity 100.0%; Pred. No. 0.00086;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17462 GCCAGATGGTCTGATCTCTGACCTGATCGGCC 17500
|||||
DB 9 GCCAGATGGTCTGATCTCTGACCTGATCGGCC 47

RESULT 12
PCT-US95-17111A-92
Sequence 92, Application PC/TUS951711A
GENERAL INFORMATION:
APPLICANT: de la Monte, Suzanne
APPLICANT: Mands, Jack R.
TITLE OF INVENTION: Neural Thread Protein Gene Expression and
NUMBER OF SEQUENCES: 121
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/1711A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/340,426
FILING DATE: 14-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609.3840002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 92:
SEQUENCE CHARACTERISTICS:
LENGTH: 85 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
PCT-US95-1711A-92

Query Match 0.1%; Score 39; DB 6; Length 85;
Best Local Similarity 100.0%; Pred. No. 0.00086;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17462 GCCAGATGGTCTGATCTCTGACCTGATCGGCC 17500
|||||
DB 9 GCCAGATGGTCTGATCTCTGACCTGATCGGCC 47

RESULT 13
US-08-222-177A-175/C
Sequence 175, Application US/08222177A
Patent No. 5582979
GENERAL INFORMATION:
APPLICANT: Weber, James L.
TITLE OF INVENTION: LENGTH POLYMORPHISMS IN
NUMBER OF SEQUENCES: 460
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dewitt Ross & Stevens, S.C.
STREET: 8000 Excelsior Drive, Suite 401
CITY: Madison
STATE: Wisconsin
COUNTRY: USA

ZIP: 53717-1914
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/222,177A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/341,562
FILING DATE: 21-APR-1989
ATTORNEY/AGENT INFORMATION:
NAME: Sara, Charles S.
REGISTRATION NUMBER: 30,492
REFERENCE/DOCKET NUMBER: 09865, 601
TELECOMMUNICATION INFORMATION:
TELEPHONE: (608) 831-2100
TELEFAX: (608) 831-2106
TELEX:
INFORMATION FOR SEQ ID NO: 175:
SEQUENCE CHARACTERISTICS:
LENGTH: 40 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
IMMEDIATE SOURCE:
CLONE: mfd42is
US-08-222-177A-175

```
Query Match Similarity      0.1% : Score 36; DB 1; length 40;  
Best Local Similarity       100.0% : Pred. No. 0.0086;  
Matches    36; Conservative   0; Mismatches     0; Indels      0; Gaps      0;
```

RESULT 14
 US-08-222-177A-71/C
 Sequence 71, Application US/08222177A
 Patent No. 3582979
 GENERAL INFORMATION:
 APPLICANT: Weber, James L.
 TITLE OF INVENTION: LENGTH POLYMORPHISMS IN
 TITLE OF INVENTION: (dc-da)n.(dc-dn) SEQUENCES AND METHODS OF USING SAME
 NUMBER OF SEQUENCES: 460
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: DEWITT ROSS & STEVENS, S.C.
 STREET: 8000 Excelsior Drive, Suite 401
 CITY: Madison
 STATE: Wisconsin
 COUNTRY: USA
 ZIP: 53717-1914
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/222.177A
 FILING DATE:
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/341,562
 FILING DATE: 21-APR-1989
 ATTORNEY/AGENT INFORMATION:
 NAME: Sara, Charles S.
 REGISTRATION NUMBER: 30,492

```

1: REFERENCE/DOCKET NUMBER: 09865,607
2: TELECOMMUNICATION INFORMATION:
3: TELEPHONE: (608) 831-2100
4:
5: TELEFAX: (608) 831-2106
6:
7:
8: TELEFAX:
9:
10: INFORMATION FOR SEQ ID NO: 71:
11:
12: SEQUENCE CHARACTERISTICS:
13:
14: LENGTH: 46 base pairs
15:
16: TYPE: nucleic acid
17:
18: STRANDEDNESS: double
19:
20: TOPOLOGY: linear
21:
22: MOLECULE TYPE: DNA (genomic)
23:
24:
25: US-08-222-177A-71

```

Query Match	0.1%;	Score 36;	DB 1;	Length 46;
Best Local Similarity	100.0%;	Pred. No. 0.0084;		
Matches	36;	Conservative	0;	Mismatches 0;
				Indels 0;
				Gaps 0;
QY	28999	TGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG	29034	
Db	44	TGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG	9	

RESULT 15
 US-08-469-802B-27/c
 Sequence 27, Application US/08469802B
 Patent No. 5741645
 GENERAL INFORMATION:
 APPLICANT: Orr, Harry T.
 APPLICANT: Rannu, Laura P.W.
 APPLICANT: Chung, Ming-Yi
 APPLICANT: Zoghbi, Huda Y.
 TITLE OF INVENTION: Gene Sequence for Spinocerebellar Ataxia
 Patent No. 5741645
 TITLE OF INVENTION: Type 1 and Method for Diagnosis
 NUMBER OF SEQUENCES: 47
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Muelling, Raasch, Gebhardt & Schwappach, P.A.
 STREET: 119 No. 5741645th Fourth Street, Suite 203
 CITY: Minneapolis

```

?      COMPUTER READABLE FORM:
?      MEDIUM TYPE: Floppy disk
?      COMPUTER: IBM PC compatible
?      OPERATING SYSTEM: PC-DOS/MS-DOS
?      SOFTWARE: Patent In Review #1.0, Version #1.25
?      CURRENT APPLICATION DATA:
?      APPLICATION NUMBER: US/08/469,802B
?      FILING DATE: 06-JUN-1995
?      CLASSIFICATION: 435
?      ATTORNEY/AGENT INFORMATION:
?      NAME: Muelling/ Ann M.
?      REGISTRATION NUMBER: 33,977
?      REFERENCE/DOCKET NUMBER: 110.00030101
?      TELECOMMUNICATION INFORMATION:
?      TELEPHONE: 612-305-1217
?      TELEFAX: 612-305-1225
?      INFORMATION FOR SEQ ID NO: 27:
?      SEQUENCE CHARACTERISTICS:
?      LENGTH: 54 base pairs
?      TYPE: nucleic acid
?      STRANDEDNESS: single
?      TOPOLOGY: linear
?      MOLECULE TYPE: DNA
?      US-08-469-802B-27

```

Mon Apr 3 08:24:20 2000

us-08-852-495c-1_copy_206000_235033.rni

Page 7

[illegible]

Search completed: April 1, 2000, 12:41:53
Job time: 291175 sec

C	2	59	0.2	103	104	A0535244	A0535244 RPCI-11-3
C	3	52	0.2	72	62	A1914923	A1914923 trt7e08.x
C	4	48	0.2	102	81	B62983	B62983 CIT-HSP-659
C	5	48	0.2	103	81	B46914	B46914 RPCI11-4A12
C	6	47	0.2	91	50	F29274	F29274 HSPD19063 H
C	7	47	0.2	105	21	T94466	T94466 yec35b02.r1
C	8	44	0.2	70	25	N84707	N84707 J0579F.Huma
C	9	44	0.2	80	50	F24490	F24490 HSPD10834 H
C	10	43	0.1	87	20	D20989	D20989 HUMGSO1971
C	11	43	0.1	101	39	A4835205	A4835205 ax64h01.s
C	12	41	0.1	90	61	A1864984	A1864984 w06c11.x
C	13	41	0.1	91	38	AA780764	AA780764 ax08f12.s
C	14	41	0.1	102	30	AA226656	AA226656 nc19f09.s
C	15	41	0.1	103	81	B90619	B90619 CIT-HSP-216
C	16	40	0.1	96	22	R67088	R67088 y130h05.s1
C	17	40	0.1	105	28	AA078003	AA078003 7H12D08 C
C	18	39	0.1	89	37	AA737634	AA737634 ca52e11.s
C	19	39	0.1	100	26	W74144	W74144 z03b05.r1
C	20	39	0.1	102	34	AA481597	AA481597 aa35e11.r
C	21	39	0.1	103	104	A0582186	A0582186 RPCI-11-4
C	22	39	0.1	105	81	B74378	B74378 CIT-HSP-202
C	23	38	0.1	74	28	AA078709	AA078709 7T01E11 C
C	24	38	0.1	101	35	AA555145	AA555145 n107g04.s
C	25	38	0.1	104	36	AA614379	AA614379 np46c03.s
C	26	37	0.1	72	45	A1349123	A1349123 ta51e08.x
C	27	37	0.1	95	34	AA457423	AA457423 aa86b02.r
C	28	36	0.1	52	20	D25845	D25845 HUMGSO4217
C	29	36	0.1	84	50	F34634	F34634 HSPD29754 H
C	30	36	0.1	101	100	AQ280224	AQ280224 CITBI-E1-
C	31	36	0.1	102	81	B80126	B80126 CIT-HSP-204
C	32	36	0.1	102	100	AQ319270	AQ319270 RPCI11-98
C	33	35	0.1	37	22	R70733	R70733 y146d12.r1
C	34	35	0.1	54	36	AA601314	AA601314 n015f06.s
C	35	35	0.1	84	34	AA501753	AA501753 ng05e01.s
C	36	35	0.1	88	38	AA809831	AA809831 ca40f11.s
C	37	35	0.1	93	21	R09732	R09732 y127a12.s1
C	38	35	0.1	94	23	R93021	R93021 yq11g03.s1
C	39	35	0.1	95	24	H61099	H61099 yf51c12.r1
C	40	35	0.1	95	35	AA578401	AA578401 n153c01.s
C	41	35	0.1	101	21	R17033	R17033 yf45d06.r2
C	42	35	0.1	101	33	AA381369	AA381369 EST94442
C	43	35	0.1	101	34	AA494273	AA494273 ng96b11.s
C	44	35	0.1	101	100	AQ278892	AQ278892 CITBI-E1-
C	45	35	0.1	103	30	AA228795	AA228795 nc14e07.s

ALIGNMENTS

RESULT 1
 AA082835 59 bp mRNA EST 21-OCT-1996
 LOCUS zn21g12.s1 Strataene neuroepithelium NT2RAMI 937234 Homo sapiens
 DEFINITION cDNA clone IMAGE:348134 3' similar to contains Alu repetitive element; mRNA sequence.
 ACCESSION AA082835
 VERSION AA082835.1 GI:1624910
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 1 (bases 1 to 59)
 Hillier, L., Lennon, G., Becker, M., Donald, M.F., Chappell, B., Chissoe, S., Dietrich, N., Dubuque, T., Favello, A., Gish, W., Hawkins, M., Hultman, M., Kucaba, T., Lacy, M., Le, N., Lewis, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L., Rohlfing, T., Scheinberg, K., Soares, M.B., Tan, F., Thirry-Meg, J., Trevisan, E., Underwood, K., Wohlmann, P., Waterston, R., Wilson, R. and Marra, M.
 Generation and analysis of 280,000 human expressed sequence tags
 JOURNAL Genome Res. 6 (9), 807-828 (1996)
 MEDLINE 97044478

COMMENT

On Sep 12, 1996 this sequence version replaced gi:1404580.
 Contact: Willson RK
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@wustl.edu
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.lnl.gov) for further information.
 Seq primer: -40M3 fwd. from Amerham.
 Location/Qualifiers
 1..59
 /organism="Homo sapiens"
 /db_xref="GDB:392650"
 /db_xref="taxon:9606"
 /clone="IMAGE:548134"
 /clone_lib="Strataene neuroepithelium NT2RAMI 937234"
 /dev_stage="Ntera-2/RA+MI neuroepithelial cells"
 /lab_host="SOLR (kanamycin resistant)"
 /note="Vector: pBluescript SK-; Site1: EcoRI; Site2: XhoI; cloned unidirectionally. Primer: Oligo dT. NM2 (Ntera-2/cl.D1) precursor cells induced with Retinoid Acid for 1 week, followed by 3 weeks in mitotic inhibitors (Replate #2). Average insert size: 1.1 kb; Uni-ZAP XR vector; -5' adaptor sequence: 5' GATTCGCGCAG 3' -3' adaptor sequence: 5' CTCGATTTTCTTTTCTTTT 3'"

FEATURES

source

BASE COUNT

11 a 14 c 16 g 18 t

Query Match

0.2%; Score 59; DB 28; Length 59;

Best Local Similarity 100.0%; Pred. No. 2.1e-08;

Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17433 TTATAGTACAGCGGGTTTCCATGTTAGCCAGAGGCTCTGATCTCTGACCTCGT 17491

Db 1 TTATAGTACAGCGGGTTTCCATGTTAGCCAGAGGCTCTGATCTCTGACCTCGT 59

RESULT 2
 A0535244/c 103 bp DNA GSS 18-MAY-1999
 LOCUS RPCI-11-317H22.TV RPCI-11 Homo sapiens genomic clone
 DEFINITION RPCI-11-317H22, genomic survey sequence.
 ACCESSION A0535244
 VERSION A0535244.1 GI:4846934
 KEYWORDS GSS.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 1 (bases 1 to 103)
 Zhao, S., Adams, M.D., Nierman, W., Malek, J., de Jong, P. and Venter, J.C.
 Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready Map Building
 Unpublished (1997)
 Contact: Shaying Zhao, William Nierman, Mark Adams
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: hbeet@igf.org
 Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Plietier de Jong (plietier@jg.med.bu.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.bu.edu/ordering) or from Research Genet cs (info@resgen.com). BAC end search page: http://www.tigr.org/cdb/humgen/bac_end_search/bac_end_search.html.
 Seq primer: 17
 Class: BAC ends.
 Location/Qualifiers

FEATURES


```
source
1. 103
/organism="Homo sapiens"
/db_xref="GDB:7621533"
/db_xref="taxon:9606"
/clone="RPC1-11-317H22"
/clone_11b="RPC1-11"
/sex="Male"
/cell_type="Lymphocytes"
/note="Vector: pBAC3.6; Site_1: EcoRI; Site_2: EcoRI;
RPC11 Human Male BAC Library"

BASE COUNT      31 a      27 c      27 g      18 t

Query Match      0.2%; Score 59; DB 104; Length 103;
Best Local Similarity 100.0%; Pred. No. 1.5e-08;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21703 AGAGACGGGGTTTACCATGTTGGCCAGCTGCTCGAAGCTCCTGACCTCAAGTATC 21761
|||||
Db 87 AGAGACGGGGTTTACCATGTTGGCCAGCTGCTCGAAGCTCCTGACCTCAAGTATC 29

RESULT 3
AI914923 72 bp mRNA EST 28-JUL-1999
LOCUS tr27e08.x1 NCI-CGAP_Ov23 Homo sapiens cDNA clone IMAGE:2219558 3'
DEFINITION similar to contains Alu repetitive element; mRNA sequence.
ACCESSION AI914923
VERSION AI914923.1 GI:5634778
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 72)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
On May 1, 1997 this sequence version replaced gi:2059395.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Christopher Koskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
www.bio.lnl.gov/db/ftp/image/image.html

JOURNAL
COMMENT
FEATURES
source
Seq primer: -40up from Gibco.
Location/Qualifiers
1. 72
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2219558"
/clone_11b="NCI-CGAP_Ov23"
/tissue_type="tumor, 5 pooled (see description)"
/lab_host="DH10B"
/note="Organ: ovary; Vector: pCMV-SPORT6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.35 kb. Tumor types include: mixed
Mullerian tumor, papillary serous, clear cell, spindle
cell. All are primary tumors, metastasis positive. Life
Technologies catalog #: 11534-013"

BASE COUNT      8 a      11 c      15 g      38 t

Query Match      0.2%; Score 52; DB 62; Length 72;

source
1. 102
/organism="Homo sapiens"
/db_xref="GDB:5491647"
/db_xref="taxon:9606"
/clone="669F15"
/clone_11b="CIR-HSP"
/sex="Male"
/cell_type="Sperm"
/note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
HindIII"

BASE COUNT      30 a      30 c      21 g      21 t

Query Match      0.2%; Score 48; DB 81; Length 102;
Best Local Similarity 100.0%; Pred. No. 4.6e-05;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1313 TTTTATATTTAGTAGAGATGGGTTCCACCATGTTGCCACGAGATG 1360
|||||
Db 59 TTTTATATTTAGTAGAGATGGGTTCCACCATGTTGCCACGAGATG 12

RESULT 4
B62983/c 102 bp DNA GSS 21-JUN-1998
LOCUS CIT-HSP-669F15.TP CIR-HSP Homo sapiens genomic clone 669F15,
DEFINITION genomic survey sequence.
ACCESSION B62983
VERSION B62983.1 GI:2636892
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 102)
determination of clone end sequences of human Bacterial Artificial
Chromosomes
Unpublished (1997)
Contact: Ung-Jin Kim
Caltech Genome Research Lab
California Institute of Technology
Division of Biology, MS 147-75, Pasadena, CA 91125, USA
Tel: 626 796 7066
Fax: 626 395 4901
Email: ung@ash.tree.caltech.edu
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.ligr.org/cdb/hungen/bac_end_search/bac_end_search.html
Seq primer: SP6
Class: BAC ends.
Location/Qualifiers
1. 102
/organism="Homo sapiens"
/db_xref="GDB:5491647"
/db_xref="taxon:9606"
/clone="669F15"
/clone_11b="CIR-HSP"
/sex="Male"
/cell_type="Sperm"
/note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
HindIII"

BASE COUNT      30 a      30 c      21 g      21 t

Query Match      0.2%; Score 48; DB 81; Length 102;
Best Local Similarity 100.0%; Pred. No. 4.6e-05;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1313 TTTTATATTTAGTAGAGATGGGTTCCACCATGTTGCCACGAGATG 1360
|||||
Db 59 TTTTATATTTAGTAGAGATGGGTTCCACCATGTTGCCACGAGATG 12

RESULT 5
B48914/c 103 bp DNA GSS 08-APR-1999
LOCUS RPC11-4A12.TP RPC1-11 Homo sapiens genomic clone RPC1-11-4A12,
DEFINITION genomic survey sequence.
ACCESSION B48914
VERSION B48914.1 GI:2601151
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
```

REFERENCE 1 (bases 1 to 103)
 AUTHORS Adams, M.D., Rounsley, S.D., Field, C.E., Bass, S., Linher, K., Golden, K., Berry, K., Granger, D., Suh, E., Wible, C., de Jong, P. and Venter, J.C.
 TITLE Use of BAC End Sequences for Sequence-Ready Map Building
 JOURNAL Unpublished (1997)
 COMMENT Contact: Mark Adams
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: madams@tigr.org
 Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieter@jlong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from Research Genetics (info@resgen.com). BAC end search page: http://www.tigr.org/tldb/hungen/bac_end_search/bac_end_search.html
 Seq primer: SP6
 Class: BAC ends.

FEATURES
 Source location/Qualifiers
 1..103
 /organism="Homo sapiens"
 /db_xref="GDB:7501163"
 /db_xref="taxon:9606"
 /clone="RPCI-11-4A12"
 /clone_1lb="RPCI-11"
 /sex="Male"
 /cell_type="Lymphocytes"
 /note="Vector: pBAC3.6; Site_1: EcoRI; Site_2: EcoRI; RPCI11 Human Male BAC Library"

BASE COUNT 30 a 28 c 30 g 15 t

ORIGIN

Query Match 0.2%; Score 48; DB 81; Length 103;
 Best Local Similarity 100.0%; Pred. No. 4.6e-05;
 Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 8728 ATCTCTGACCTCGATCGCCGCCCTCGCTCCCAAGCTGGG 8775
 ||||||||||||||||||||||||||||||||||||||||||||
 Db 53 ATCTCTGACCTCGATCGCCGCCCTCGCTCCCAAGCTGGG 6

RESULT 6
 F29274/c 91 bp mRNA EST 13-MAY-1999
 LOCUS HSPD19063 HM3 Homo sapiens cDNA clone s4000074G04, mRNA sequence.
 DEFINITION F29274
 ACCESSION F29274
 VERSION F29274.1 GI:4814900
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 91)
 AUTHORS Lanfranchi, G., Muraro, T., Caldara, F., Pacchioni, B., Pallavicini, A., Pandolfo, D., Toppo, S., Trevisan, S., Scarso, S. and Valle, G.
 TITLE Identification of 4370 expressed sequence tags from a 3'-end-specific cDNA library of human skeletal muscle by DNA sequencing and filter hybridization
 JOURNAL Genome Res. 6 (1), 35-42 (1996)
 MEDLINE 96276048
 COMMENT On Apr 7, 1998 this sequence version replaced gi:3034419.
 CONTACT: Valle G.
 CRI Biotechnology Centre
 University of Padua
 Via Trieste 75, 35121 Padua, Italy
 ABI Chromatograms and other information are available on WWW at http://grip.bio.unipd.it.
 FEATURES
 Source location/Qualifiers
 1..91

REFERENCE 1 (bases 1 to 103)
 AUTHORS Adams, M.D., Rounsley, S.D., Field, C.E., Bass, S., Linher, K., Golden, K., Berry, K., Granger, D., Suh, E., Wible, C., de Jong, P. and Venter, J.C.
 TITLE Use of BAC End Sequences for Sequence-Ready Map Building
 JOURNAL Unpublished (1997)
 COMMENT Contact: Mark Adams
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: madams@tigr.org
 Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieter@jlong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from Research Genetics (info@resgen.com). BAC end search page: http://www.tigr.org/tldb/hungen/bac_end_search/bac_end_search.html
 Seq primer: SP6
 Class: BAC ends.

FEATURES
 Source location/Qualifiers
 1..103
 /organism="Homo sapiens"
 /db_xref="GDB:7501163"
 /db_xref="taxon:9606"
 /clone="sa4000074G04"
 /clone_1lb="HM3"
 /sex="Female"
 /tissue_type="pectoral muscle (after mastectomy)"
 /note="Vector: pCDNAII (Invitrogen); Site_1: BstXI; Site_2: NotI; The library is not subtracted nor normalized. The first strand cDNA was primed with a biotinylated oligo-dT-NotI primer (5'-biotin-AACCCGCGCTCGAGCGGCCCTTTT-3'). The ds cDNA was sonicated and size-selected in the range 350-550 bp. The 3' specific fragments were selected by streptavidin coated magnetic beads, ligated to non palindromic BstXI adapters. NotI digested and directionally cloned into BstXI-NotI cut pCDNAII vector."

BASE COUNT 18 a 19 c 38 g 16 t

ORIGIN

Query Match 0.2%; Score 47; DB 50; Length 91;
 Best Local Similarity 100.0%; Pred. No. 0.0001;
 Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 8621 TCCTGCTCAGCCTCCGAGTAGCTGGAGCTACAGCGCCCTGCACG 8667
 ||||||||||||||||||||||||||||||||||||||||||||
 Db 63 TCCTGCTCAGCCTCCGAGTAGCTGGAGCTACAGCGCCCTGCACG 17

RESULT 7
 T94466/c 105 bp mRNA EST 24-MAR-1995
 LOCUS Y63502.1 Stragene lung (#937210) Homo sapiens cDNA clone IMAGE:119691 5' similar to contains Alu repetitive element; mRNA sequence.
 ACCESSION T94466
 VERSION T94466.1 GI:727954
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 105)
 AUTHORS Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chiappelli, B., Chisoe, S., Dietrich, N., Dubuque, T., Favell, A., Gish, W., Harkins, M., Hultman, M., Kucaba, T., Lacy, M., Le, N., Mardis, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L., Rohlfing, T., Schellenberg, K., Soares, M.B., Tan, F., Thierly-Meg, J., Trevaaskis, E., Underwood, K., Wohldmann, P., Waterston, R., Wilson, R. and Marra, M.
 TITLE Generation and analysis of 280,000 human expressed sequence tags
 JOURNAL Genome Res. 6 (3), 807-828 (1996)
 MEDLINE 97044478
 COMMENT Contact: Wilson R.
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 Insert Size: 614
 High quality sequence stops: 70 Source: IMAGE Consortium, LNL This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 Insert Length: 614 Std Error: 0.00
 Seq primer: M13RPI
 High quality sequence stop: 70.
 FEATURES
 Source location/Qualifiers
 1..105
 /organism="Homo sapiens"
 /db_xref="GDB:487980"
 /db_xref="taxon:9606"
 /clone="IMAGE:119691"

/clone_1lb="Stratagene lung (#937210)"
/sex="male"
/dev_stage="72 years"
/lab_host="SOLR cells (Kanamycin resistant)"
/note="Organ: lung; Vector: pBluescript SK-; Site.1:
EcoRI; Site.2: XhoI; Cloned unidirectionally. Primer:
0.11g0 dt. normal lung. Average insert size: 1.0 kb;
Uni-ZAP XR Vector; ~5' adaptor sequence: 5' GAATTCGGCAGCAG
3' ~3' adaptor sequence: 5' CTCGAGTTTCTTTTCTTTT 3' "

BASE COUNT
ORIGIN

31 a 20 c 34 g 18 t 2 others

Query Match 0.2%; Score 47; DB 21; Length 105;
Best Local Similarity 100.0%; Pred. No. 9.4e-05;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25288 CCTCTGGGTTCAAGTATCTCTGCTCAGCCTCTGAGTACGTG 25334
DB 88 CCTCTGGGTTCAAGTATCTCTGCTCAGCCTCTGAGTACGTG 42

RESULT 8
LOCUS N84707 70 bp mRNA EST 01-APR-1996
DEFINITION J0579F Human fetal heart, Lambda ZAP Express Homo sapiens cDNA
clone J0579 5' similar to REPETITIVE ELEMENT ALU, mRNA sequence.
ACCESSION N84707
VERSION N84707.1 GI:1260332
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 70)
Liew.C.C.
CDNAs from fetal heart (1996)
Unpublished (1996)
Contact: Liew CC
Department of Laboratory Medicine and Pathobiology
University of Toronto
Banting Institute, 100 College St., Toronto, Ontario, M5G1L5
Tel: 4169788758
Fax: 4169785650
Email: liewcc@utcc.utoronto.ca
Seq primer: GAATTAACCTCTCCTAAAGG.
Location/Qualifiers

FEATURES
source
1. 70
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="J0579"
/clone_1lb="human fetal heart, Lambda ZAP Express"
/lab_host="E. coli XL1-Blue"
/note="Vector: Lambda ZAP Express; Site.1: EcoRI; Site.2:
XhoI; mRNA was purified from human fetal hearts (8-10
weeks). cDNA was synthesized using a XhoI-Oligo dt
adaptor-primer. EcoRI adaptors were ligated, followed by
digestion with XhoI, for directional cloning into
predigested Lambda ZAP Express."

BASE COUNT 24 a 18 c 15 g 13 t
ORIGIN

Query Match 0.2%; Score 44; DB 25; Length 70;
Best Local Similarity 100.0%; Pred. No. 0.0011;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17445 GGGGTTTCAACATGTTAGCCAGATGCTCGATCTCTGACCT 17488
DB 50 GGGGTTTCAACATGTTAGCCAGATGCTCGATCTCTGACCT 7

RESULT 9

F24490/c
LOCUS F24490 80 bp mRNA EST 13-MAY-1999
DEFINITION HSPD10834 HM3 Homo sapiens cDNA clone s4000013A06, mRNA sequence.
ACCESSION F24490
VERSION F24490.1 GI:4810116
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 80)
Lanfranchi,G., Muraro,T., Caldara,F., Pacchioni,B., Pallavicini,A.,
Pantolfo,D., Toppo,S., Trevisan,S., Scarso,S. and Valle,G.
Identification of 4370 expressed sequence tags from a
3'-end-specific cDNA library of human skeletal muscle by DNA
sequencing and filter hybridization
Genome Res. 6 (1), 35-42 (1996)

JOURNAL MEDLINE
COMMENT On Jun 5, 1998 this sequence version replaced gi:318834.
96276048
Contact: Valle G.
CIRI Biotechnology Centre
University of Padua
Via Trieste 75, 35121 Padua, Italy
ABI Chromatograms and other information are available on WWW at
<http://grup.bio.unipd.it>.

FEATURES
source
1. 80
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="s4000013A06"
/clone_1lb="HM3"
/sex="female"
/tissue_type="pectoral muscle (after mastectomy)"
/note="Vector: pCDNMI (Invitrogen). Site.1: BstXI;
Site.2: NotI. The library is not subtracted nor normalized.
Lanfranchi. This library is not subtracted nor normalized.
The first strand cDNA was primed with a biotinylated
oligo-dT-NotI primer
(5'-biotin-AACCGGCTCGAGCGGCGGCTTTTCTTTTCTTTT-3'). The
ds cDNA was sonicated and size-selected in the range
350-550 bp. The 3' specific fragments were selected by
streptavidin coated magnetic beads, ligated to
non-palindromic BstXI adaptors, NotI digested and
directionally cloned into BstXI-NotI cut pCDNMI vector."

BASE COUNT 18 a 21 c 19 g 22 t
ORIGIN

Query Match 0.2%; Score 44; DB 50; Length 80;
Best Local Similarity 100.0%; Pred. No. 0.00098;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19018 GGCCTCCCAAGTCTGGATTACGACGATGAGCCACTGGCCCC 19061
DB 80 GGCCTCCCAAGTCTGGATTACGACGATGAGCCACTGGCCCC 37

RESULT 10
LOCUS D20989 87 bp mRNA EST 30-JUL-1996
DEFINITION HUMGS01971 Human promyelocyte Homo sapiens cDNA clone mp0383 3',
HMG sequence.
ACCESSION D20989
VERSION D20989.1 GI:504809
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 87)
Okubo,K., Fukushima,A., Yoshii,J., Niyama,T., Kojima,Y.,
Yoshinari,H., Arimoto,J. and Matsubara,K.
Gene expression of human promyelocytic cell line HL60 before and

REFERENCE
AUTHORS
TITLE

JOURNAL
COMMENT

after induction of differentiation. A new application of 3' directed cDNA sequencing (1993)
Unpublished (1993)
Contact: Okubo, K., Fukushima, A., Yoshii, J., Miyama, T., Kojima, Y., Yoshinari, H., Arimoto, J. and Matsubara, K.
Institute for Molecular and Cellular Biology
Osaka University
3-1 Yamada-Oka, Suita, Osaka 565, Japan.

FEATURES
source

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="mp0383"
/clone_lib="Human promyelocyte"
/note="Female, adult, cell_line = HL60, cell_type =
promyelocyte."
BASE COUNT      13 a      23 c      17 g      34 t
ORIGIN

```

```

Query Match      0.1% ; Score 43 ; DB 20 ; length 87 ;
Best Local Similarity 100.0% ; Pred. No. 0.0019 ;
Matches 43 ; Conservative 0 ; Mismatches 0 ; Indels 0 ; Gaps 0 ;

```

Qy 25449 CCTCAGCCTCCCAAGTGTGGATTACAGGTGTGAGCCACCA 25491

Db 11 CCTCAGCCTCCCAAAGTGTCTGGATTACAGGTGTGAGCCACCA 53

RESULT 11
AA835205

LOCUS	EST	23-FEB-1998
DEFINITION	101 bp	mRNA
	AK64401.s1	Barsted pancreas HPLR1 Homo sapiens cDNA clone
	INACG:1412689	3' similar to contains Alu repetitive
	element; contains element	KER repetitive element ; , mRNA sequence

VERSION AA835205.1 GI:2908933

SOURCE

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Mammalia,
Euthera; Primates; Catarrhini; Homidae; Homo.

AUTHORS

Krizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F., Theising, B., White, Y., Wylie, T., Waterston, R. and Wilson, R.

JOURNAL Unpublished (1997)

COMMENT

Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: estewatson.wustl.edu

This clone is available royalty-free through LNC₊; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -40m13 fwd: ET from Amerham.

FEATURES
SOURCE

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:141268"
/clone_id="Barstlead pancreas HPLRb1"
/sex="female"
/dev_stage="adult, 34 years"
/lab_host="DH10B"
/note="Organ: pancreas; Vector: pT7n3D-Pac (pharmacia)
with a modified polylinker. Site1: EcoRI; Site2: NotI;
1st strand cDNA was primed with a Not I - oligo(dT) primer"
15'
IGTTCACATCTGAAAGTGAGGCGCGCCCTTTTTTTTTTTTTTTTTTTTTT

```

BASE COUNT
ORIGIN

3'] double-stranded cDNA was ligated to Eco RI adaptors
[ATCGGACGCTTG], digested with Not I and cloned into the
Not I and Eco RI sites of the modified pRT3 vector.
Library constructed by Bob Barstead. "

```

Query Match: 0.1%; score 43; DB 39; length 101;
Best Local Similarity 100.0%; Pred. No. 0.0018;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

8577 ATCTGGCTCACTGCAAGCTCCGGCTCCCGGGTTCACGCCATT 8619
OY

Db 48 ATCTCGGCTCACTGCAAGCTCGGCTCCCGGGTTACGCCATT 90

RESULT 12
AI864984

LOCUS	A1864984	90 bp	mRNA	EST	30-AUG-1999
DEFINITION	w06c11.x1 NCI-CGAP_Lym12 Homo sapiens cDNA clone IMAGE:2411540 3 similar to contains Alu repetitive element;; mRNA sequence.				
ACCESSION	A1864984				
VERSION	A1864984.1	GI:5529091			

SOURCE human.

REFERENCE
AUTHORS
TITLE

Eutheria; Primates; Carnivora; Hominoidea; Homo.
1 (bases 1 to 90)
NCI-CCGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Informatics (NCI)

JOURNAL Unpublished (1997)

Contact: Robert Strausberg, Ph.D.
Tel.: (201) 406-1550

Email: Robert_Strau

DNA Sequencing by: Washington University Genome Sequencing Center

www-bio.11n1.gov/bbrp/image/image.html

FEATURES

Seq primer: -40UP from Gibco.

source

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_1b="NCI CGAP_Lym12"
/tissue_type="Lymphoma, follicular mixed small and large

```

```
/lab_host="DH10B"
```

SALT; Site_2: NotI; Cloned unidirectionally. Primer:

catalog #: 11547-015"

ORIGIN

Query Match	0.13;	Score 41;	DB 61;	Length 90;
Best Local Similarity	100.08;	Pred. No. 0.0081;		
Match 41; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0

8577 ATCTGGCTCACTGCAAGCTCCGCTCCCGGGTTACGCCA 8617
QY

Db 50 ATCTCGGCTACTGCAAGCTCCGGCTCCGGGTTACGCCA 90

RESULT 13

AA780764

LOCUS AA780764 91 bp mRNA EST 05-FEB-1998
 DEFINITION AC68f12.s1 Stratagene fetal retina 937202 Homo sapiens cDNA clone
 IMAGE:667791 3' similar to contains Alu repetitive element; mRNA
 sequence.
 ACCESSION AA780764
 VERSION AA780764.1 GI:2840095
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE 1 (bases 1 to 91)
 Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisler, G., Jost, S.,
 Krizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M.,
 Martin, J., Moore, B., Schellenger, K., Stepien, M., Tan, F.,
 Thelshing, B., White, Y., Wylie, T., Waterston, R. and Wilson, R.
 WashU-NCI human EST Project
 Unpublished (1997)
 TITLE On Sep 12, 1996 this sequence version replaced gi:1407381.
 JOURNAL Contact: Wilson RK
 COMMENT Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@wustl.edu
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.lnl.gov) for further information.
 Seq primer: 40m13 fwd. ET from Amerisham.
 Location/Qualifiers
 1..91
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:667791"
 /clone_1lb="Stratagene fetal retina 937202"
 /sex="mixed"
 /lab_host="SOLR (Kanamycin resistant)"
 /note="Vector: pBluescript SK-; Site.1: EcoRI; Site.2:
 XhoI; Cloned unidirectionally. Primer: oligo dT. Pooled
 retinal tissue. Average insert size: 1.0 kb; Uni-ZAP XR
 Vector: -5' adaptor sequence: 5' GATTTCGGACGAG 3' -3'
 adaptor sequence: 5' CTCGAGTTTCTTTTCTTTT 3' -3'
 BASE COUNT 15 a 33 c 24 g 19 t
 ORIGIN
 Query Match 0.1%; Score 41; DB 38; Length 91;
 Best Local Similarity 100.0%; Pred. No. 0.0081;
 Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 24778 CAATCGGCTCAGTCAACCTCCGCTCCGGGTCACAGC 24818
 Db 43 CAATCGGCTCAGTCAACCTCCGCTCCGGGTCACAGC 83
 RESULT 14
 LOCUS AA226656 102 bp mRNA EST 21-AUG-1997
 DEFINITION nc19109.s1 NCI-CGAP_Prl Homo sapiens cDNA clone IMAGE:1008617
 similar to contains Alu repetitive element; mRNA sequence.
 ACCESSION AA226656
 VERSION AA226656.1 GI:1847964
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE 1 (bases 1 to 102)
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT On Sep 12, 1996 this sequence version replaced gi:1309483.
 Contact: Robert Strausberg, Ph.D.

Tel: (301) 496-1550
 Email: Robert.Strausberg@nih.gov
 Tissue Procurement: W. Marston Linehan, M.D., Rodrigo Chuagui,
 M.D., Michael Emmert-Buck, M.D., Ph.D.
 CDNA Library Preparation: David B. Krizman, Ph.D.
 CDNA Library Arrayed by: Genome Systems Inc., Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 www.bio.lnl.gov/bdrip/image/image.html
 Insert Length: 282 Std Error: 0.00
 Seq primer: -41m13 fwd. ET from Amerisham
 High quality sequence stop: 82.
 Location/Qualifiers
 1..102
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:1008617"
 /clone_1lb="NCI-CGAP_Prl"
 /sex="Male"
 /dev_stage="45 years old"
 /lab_host="DH10B"
 /note="Vector: PAMp10; Site.1: NotI; Site.2: EcoRI; 1st
 strand cDNA was primed with oligo(dT)17 on 50 ng of
 DNase-treated, total cellular RNA obtained from
 5,000-10,000 microdissected, histologically normal
 prostate epithelial cells. Double-stranded cDNA was
 ligated to EcoRI adaptors, 5 cycles of PCR applied to the
 cDNA with an adaptor-specific primer, and the resulting
 PCR product subcloned into pAMP10 by the UDG-cloning
 method (Life Technologies). Average insert size is 600
 bp. NOTE: Not directionally cloned. This library was
 constructed by David Krizman."
 BASE COUNT 13 a 32 c 26 g 30 t 1 others
 ORIGIN
 Query Match 0.1%; Score 41; DB 30; Length 102;
 Best Local Similarity 100.0%; Pred. No. 0.0076;
 Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 18798 CTCTGTACCCAGCTGAGTGCAGTGGCGATCTCGCT 18838
 Db 21 CTCTGTACCCAGCTGAGTGCAGTGGCGATCTCGCT 61
 RESULT 15
 LOCUS B90619/c 103 bp DNA GSS 25-JUN-1998
 DEFINITION CIT-HSP-2163G1.TR CIT-HSP Homo sapiens genomic clone 2163G1,
 genomic survey sequence.
 ACCESSION B90619
 VERSION B90619.1 GI:2973099
 KEYWORDS GSS.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE 1 (bases 1 to 103)
 Adams, M.D., Rounsley, S.D., Zhao, S., Field, C.E., Bass, S., Linher, K.,
 Golden, K., Berry, K., Granger, D., Suh, E., Wible, C., Shizuya, H.,
 Simon, M. and Venter, J.C.
 Use of a random BAC End sequence Database for Sequence-Ready Map
 Building (1998)
 TITLE Unpublished (1998)
 JOURNAL Other-GSS: CIT-HSP-2163G1.TF
 COMMENT Contact: Mark Adams
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0200
 Fax: 301 838 0208

Email: mtdams@ligr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page: http://www.ligr.org/ldb/hungen/bac_end_search/bac_end_search.html
Seq primer: M13 Reverse
Class: BAC ends.

FEATURES
source location/Qualifiers

1. 103
/organism="Homo sapiens"
/db_xref="GDB:7100363"
/db_xref="taxon:9606"
/clone="216361"
/clone_id="CIT-HSP"
/sex="Male"
/cell_type="Sperm"
/note="Vector: pbelBAC11; Site_1: HindIII; Site_2:
HindIII"

BASE COUNT 27 a 31 c 27 g 18 t
ORIGIN

Query Match 0.18; Score 41; DB 81; Length 103;
Best Local Similarity 100.0%; Pred.No.0.0075;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 17507 GCCTCCCAAGTGTGAGATTACAGGTGTGAGCCACGCGC 17547
|||||
Db 100 GCCTCCCAAGTGTGAGATTACAGGTGTGAGCCACGCGC 60

Search completed: April 1, 2000, 06:35:49
Job time: 271203 sec